ASIMOV

L'univers de la science



InterEditions

Voici, pour la première fois en français, la quatrième édition d'un best-seller devenu depuis longtemps outre-Atlantique une référence obligatoire pour tout adulte cultivé, tout adolescent curieux de science, tout scientifique soucieux de se tenir au courant des derniers progrès dans sa spécialité.

L'auteur couvre tous les domaines des sciences physiques et biologiques, rendant compte des tout derniers développements dans les diverses disciplines (trous noirs, quarks, quasars, informatique, intelligence artificielle, robotique, progrès dans les domaines de la médecine, de l'astrophysique, etc.) mais adoptant également une perspective historique. Il nous fait ainsi partager la passion des scientifiques de tous les temps et de toutes les cultures et les péripéties de leurs découvertes.

Réunie en un seul volume, c'est tout simplement la somme des connaissances scientifiques conquises par l'humanité, de l'Antiquité à nos jours, que nous offre ce surprenant auteur.

Le style d'Isaac Asimov est clair, et tous s'accordent à lui ▸

L'UNIVERS DE LA SCIENCE

Isaac Asimov

L'UNIVERS DE LA SCIENCE

VERSION FRANÇAISE

Françoise Balibar, Claude Guthmann, Alain Laverne et Jean-Pierre Maury.

Christine Coulondre, Marie-Jo Masse et John Reams.



InterEditions

L'édition originale de cet ouvrage a été publiée aux États-Unis par Basic Books, Inc., Publishers, New York sous le titre *Asimov's New Guide to Science*. © 1984 by Basic Books, Inc.

ISBN 2-7296-0141-4

Pour

Janet Jeppson Asimov

qui partage mon intérêt pour la science

et tous les autres aspects de ma vie

Table des matières

PREMIÈRE PARTIE
Les sciences physiques

SECONDE PARTIE
Les sciences biologiques

Planches

L'UNIVERS
DE
LA SCIENCE

Chapitre 1

Qu'est-ce que la science ?

Au commencement (ou presque) était la curiosité.

La curiosité, le désir impérieux de savoir, n'est pas une caractéristique de la matière inerte. Cela ne semble pas non plus caractériser certaines formes de vie, que pour cette raison nous avons du mal à considérer comme vraiment vivantes.

Un arbre ne manifeste, à l'égard de son environnement, aucune curiosité que nous puissions déceler ; une éponge ou une huître pas davantage. Le vent, la pluie, les courants marins leur amènent ce qui leur est nécessaire, et ils en prennent ce qu'ils peuvent. Si le hasard leur apporte le feu, le poison, les prédateurs ou les parasites, ils meurent aussi stoïquement et discrètement qu'ils ont vécu.

Mais très tôt dans le développement de la vie, certains organismes ont acquis la liberté de se déplacer. Ceci entraînait un progrès considérable dans leur contrôle de l'environnement. Un organisme mobile n'est plus obligé d'attendre, dans un calme inébranlable, que la nourriture vienne vers lui ; il part à sa recherche.

C'est ainsi que le monde connut l'aventure – et la curiosité. L'individu qui hésitait dans la compétition pour la nourriture, qui était trop timide dans ses recherches, mourait de faim. Très tôt, la curiosité à l'égard de l'environnement était devenue la condition de la survie.

La paramécie unicellulaire, quand elle cherche autour d'elle, ne peut pas avoir de volontés et de désirs conscients comme les nôtres, mais elle a une tendance, même si c'est « seulement » une tendance physico-chimique, qui la fait se comporter comme si elle explorait les environs à la recherche de nourriture, d'un abri, ou des deux. Et cet « acte de curiosité » est ce que nous reconnaissons le plus facilement comme inséparable du type de vie qui nous est le plus proche.

A mesure que les organismes devenaient plus perfectionnés, leurs organes sensoriels se multipliaient et devenaient plus complexes et plus délicats. Des messages plus nombreux et plus variés parvenaient de l'environnement. En même temps (cause ou conséquence, nous ne saurions le dire) le système nerveux, appareil vivant chargé d'interpréter et d'enregistrer les informations fournies par les organes des sens, devenait de plus en plus complexe.

LE DÉSIR DE SAVOIR

Il arrive un moment où la capacité de recevoir, d'enregistrer et d'interpréter les messages de l'extérieur dépasse les besoins élémentaires. Un organisme peut être rassasié, et momentanément à l'abri de tout danger. Que va-t-il faire ?

Il pourrait retomber dans la torpeur de l'huître. Mais les organismes du niveau le plus élevé, au moins, manifestent encore un puissant instinct d'exploration. Curiosité sans but, pourrait-on dire. Et pourtant, même si nous nous en moquons, c'est elle qui nous sert à évaluer l'intelligence. Le chien, dans ses moments de loisir, flaire çà et là, dresse les oreilles à des bruits qui nous échappent ; aussi, nous l'estimons plus intelligent que le chat, qui occupe ses loisirs à une toilette minutieuse, ou s'étire voluptueusement avant de s'endormir. Plus le cerveau est perfectionné, plus fort est le goût de l'exploration, plus grand est l'« excès de curiosité ». La curiosité du singe est proverbiale. Son petit cerveau remuant doit travailler, et travaille sur tout ce qui se présente à lui. Et sous ce rapport, comme sous beaucoup d'autres, l'homme est un supersinge.

Le cerveau humain est le morceau de matière le plus merveilleusement organisé de l'Univers connu, et sa capacité de recevoir, de classer et d'enregistrer l'information dépasse très largement les besoins ordinaires de la vie. On a estimé que, au cours de sa vie, un être humain pouvait emmagasiner 15 000 milliards d'informations.

C'est à cette disponibilité que nous devons d'être vulnérables à une maladie très pénible, l'ennui. Un être humain, placé dans une situation où il n'a aucune occasion d'utiliser son cerveau si ce n'est pour assurer sa survie, ressentira une quantité croissante de symptômes inquiétants, jusqu'à connaître de graves troubles mentaux. Le fait est que l'être humain normal possède une curiosité intense et irrésistible. S'il n'a pas la possibilité de la satisfaire d'une façon immédiatement utile, il la satisfera par d'autres moyens – y compris des moyens regrettables, qui sont à l'origine d'expressions courantes comme « mêlez-vous de vos affaires » ou « la curiosité est un vilain défaut ».

La force irrésistible de la curiosité, même entraînant les plus grands risques, se reflète dans les mythes et les légendes de l'humanité. Les Grecs racontaient l'histoire de la boîte de Pandore. Pandore, la première femme,

s'était vu offrir une boîte qu'on lui avait interdit d'ouvrir. Immédiatement, et naturellement, elle l'avait ouverte, libérant sur le monde son contenu de maux, maladies, famines, haines et quantité d'autres désastres.

Dans l'histoire biblique de la tentation d'Ève, il semble (en tout cas, il me semble) que le serpent n'a pas eu une tâche bien difficile, et qu'il aurait aussi bien pu se taire ; la curiosité d'Ève aurait suffi à la pousser à goûter le fruit défendu, sans qu'il soit besoin d'un tentateur extérieur. Si vous avez une vision allégorique de la Bible, vous pouvez considérer le serpent simplement comme une représentation de ce besoin intérieur. Dans le tableau traditionnel d'Ève debout sous l'arbre avec le fruit défendu à la main, le serpent enroulé dans les branches pourrait recevoir l'étiquette « curiosité ».

Si la curiosité, comme toutes les tendances humaines, peut être utilisée bassement – sous la forme de l'invasion de la vie privée qui a donné au mot « curiosité » ses connotations désagréables – elle n'en est pas moins l'une des qualités les plus nobles de l'esprit humain. Car sa définition la plus simple est « désir de savoir ».

Ce désir trouve ses premières satisfactions dans les réponses aux besoins pratiques de la vie de tous les jours : comment semer et récolter au mieux, comment fabriquer au mieux arcs et flèches, comment tisser le mieux possible – en somme, des questions techniques. Mais quand on domine ces techniques élémentaires, quand les besoins matériels sont satisfaits, que faire ? Inévitablement, le désir de savoir conduit alors à des activités moins limitées et plus complexes.

Il semble certain que les « beaux-arts » (destinés à satisfaire des besoins spirituels confus et infinis) sont nés d'un ennui insupportable. Bien sûr, on peut leur trouver des fonctions et des excuses. Peintures et statuettes, par exemple, servaient d'amulettes de fertilité et de symboles religieux. Mais on ne peut pas s'empêcher de supposer que les objets ont existé avant qu'on leur découvre une fonction.

Dire que les beaux-arts sont nés d'un sens du beau, ce serait peut-être aussi mettre la charrue avant les bœufs. Une fois inventés les beaux-arts, leur développement et leur raffinement dans le sens du beau était inévitable, mais de toute façon, il était obligatoire qu'ils apparaissent. A coup sûr, ils sont antérieurs à tout usage possible et à tout besoin imaginable, si ce n'est la nécessité d'occuper l'esprit aussi complètement que possible.

Ce n'est pas seulement que l'élaboration d'une œuvre d'art occupe de façon satisfaisante l'esprit de son créateur ; sa contemplation rend un service analogue au spectateur. Une grande œuvre est grande précisément parce qu'elle offre quelque chose que l'on ne trouve pas ailleurs. Elle contient suffisamment d'informations, de complexité, pour inciter le cerveau à un effort dépassant les besoins habituels ; à moins que l'on ne soit enlisé dans la routine ou complètement abruti, cet effort est plaisant.

Mais si la pratique des beaux-arts est une solution satisfaisante au problème des loisirs, elle a un inconvénient : en plus d'un esprit actif et créatif, elle exige de l'habileté manuelle. Il peut être aussi intéressant de se livrer à des activités intellectuelles qui n'exigent aucune habileté physique. Et, bien entendu, il existe de telles activités : la recherche du savoir, non pas pour s'en servir dans un but pratique, mais pour lui-même.

Ainsi, le désir de savoir semble mener à des sphères successives d'abstraction croissante, occupant de plus en plus l'esprit – du savoir-faire au savoir « faire beau », puis au savoir tout court.

C'est seulement la recherche du savoir pour lui-même qui peut conduire à des questions comme « A quelle hauteur est le ciel ? » ou « Pourquoi une pierre tombe-t-elle ? ». C'est là de la pure curiosité – la plus désintéressée et donc peut-être la plus péremptoire. Après tout, cela ne sert à rien de savoir à quelle hauteur se trouve le ciel, ou pour quelle raison la pierre tombe. Les hauteurs célestes n'interviennent pas dans la vie de tous les jours, et quant à la pierre, ce n'est pas en sachant pourquoi elle tombe qu'on l'esquivera plus facilement, ou que le choc sera moins rude si elle nous atteint. Pourtant, il y a toujours eu des gens pour se poser des questions apparemment aussi inutiles et pour chercher à y répondre – par pure curiosité, par pur besoin de faire travailler l'esprit.

La méthode évidente pour résoudre de tels problèmes consiste à fabriquer une réponse satisfaisante du point de vue esthétique, une réponse qui ait suffisamment d'analogies avec des choses déjà connues pour être compréhensible et plausible. Le mot « fabriquer » peut paraître dénué de romantisme. Les Anciens considéraient volontiers l'activité de découverte comme le résultat de l'inspiration des muses, ou comme une révélation céleste. Mais de toute façon, qu'il s'agisse de révélation, d'inspiration, ou de l'activité imaginaire qui anime l'art du conteur, les explications reposaient largement sur des analogies. La foudre détruit et terrifie, mais elle ressemble à une arme de jet, et fait les mêmes dommages qu'une arme de jet – d'une puissance fantastique. Une telle arme demande un lanceur aussi fantastique, et ainsi l'éclair devient le marteau de Thor ou le javelot éblouissant de Zeus. La superarme exige et révèle le surhomme.

Ainsi naissent les mythes. Les forces naturelles, personnifiées, deviennent des dieux. Les mythes réagissent les uns sur les autres, s'édifient et s'embellissent avec le passage des générations de conteurs, jusqu'à ce que le point de départ soit difficile à discerner. Certains mythes dégénèrent en historiettes ou en gaillardises, tandis que d'autres prennent une valeur morale qui les fait accepter par certaines religions de premier plan.

De même que les arts peuvent être « beaux » ou « appliqués », de même les mythes peuvent être préservés pour leur beauté ou appliqués aux besoins physiques des hommes. Par exemple, les premiers cultivateurs étaient directement concernés par la pluie et ses caprices. La pluie fertilisante tombant du ciel sur la terre évoquait évidemment l'acte sexuel, et en personnalisant le ciel et la terre, les hommes ont trouvé une explication facile aux pluies et aux sécheresses. La déesse Terre ou le dieu Ciel était content ou fâché, suivant le cas. Une fois ce mythe admis, les fermiers disposaient d'une base plausible pour l'art du faiseur de pluies – apaiser le dieu par des cérémonies appropriées. Ces cérémonies étaient souvent de caractère orgiaque – une tentative de prêcher par l'exemple le ciel et la terre.

LES GRECS

Les mythes grecs sont parmi les plus beaux et les plus sophistiqués de notre patrimoine littéraire et culturel occidental. Mais ce sont aussi les Grecs qui, le moment venu, ont inventé la manière opposée de regarder l'Univers – comme quelque chose d'impersonnel et d'inanimé. Pour les faiseurs de mythes, tous les aspects de la nature étaient essentiellement humains dans leur imprévisibilité. Si puissant et majestueux que pût être

le personnage, si surhumains les pouvoirs de Zeus, Ishtar, Isis, Marduk ou Odin, ces dieux étaient aussi – comme de simples mortels – frivoles, capricieux, émotifs, capables de réagir violemment à des choses insignifiantes, sensibles à des « pots de vin » puérils. Tant que l'Univers était sous le contrôle de divinités aussi fantasques et imprévisibles, il n'y avait pas d'espoir de le comprendre, seulement un espoir fugitif de se le concilier. Mais selon le nouveau point de vue des penseurs grecs, l'Univers devenait une machine, gouvernée par des lois inflexibles. Les philosophes grecs se mirent à l'exercice intellectuel passionnant de chercher à découvrir ces lois de la nature.

Le premier à le faire, suivant la tradition grecque, fut Thalès de Milet, vers 600 av. J.-C. Les écrivains grecs ultérieurs devaient lui attribuer une foule de découvertes, et c'est peut-être lui qui a apporté au monde grec la masse de connaissances accumulées par les Babyloniens : il passe pour avoir prédit, en 585, une éclipse de Soleil qui a effectivement eu lieu.

En s'attelant à ce problème, les Grecs admettaient, bien sûr, que la nature jouerait franc jeu ; que si on l'attaquait de la manière correcte, elle révélerait ses secrets, et ne changerait pas ses positions au milieu de la partie. (Plus de deux mille ans plus tard, Albert Einstein exprimait cette position sous une forme un peu différente : « Dieu peut être subtil, mais pas méchant ».) Ils pensaient aussi que les lois naturelles, une fois trouvées, seraient compréhensibles. Cet optimisme grec n'a jamais complètement abandonné l'espèce humaine.

Faisant confiance à la nature pour jouer franc jeu, les hommes devaient mettre au point une méthode pour apprendre à tirer des phénomènes les lois sous-jacentes. Progresser d'un point à un autre grâce à des règles d'argumentation bien établies, c'est « raisonner ». L'intuition peut guider la recherche, mais c'est le raisonnement qui doit servir à vérifier les théories. Pour prendre un exemple simple, sachant que le cognac à l'eau, le whisky à l'eau, la vodka à l'eau et le rhum à l'eau enivrent, on pourrait en déduire que ce qui enivre, c'est ce que toutes ces boissons ont en commun, à savoir l'eau. Il y a quelque chose qui ne va pas dans ce raisonnement, mais quoi ? Ce n'est pas immédiatement évident. Dans des cas plus subtils, l'erreur peut être vraiment très difficile à découvrir.

Le dépistage des erreurs de raisonnement a amusé tous les penseurs depuis le temps des Grecs jusqu'à nos jours. Et nous devons les premières fondations de la logique systématique à Aristote qui, au IVe siècle av. J.-C., a rassemblé le premier les règles du raisonnement rigoureux.

L'essentiel du jeu intellectuel de l'homme contre la nature tient en trois points. Premièrement, recueillir des informations sur tel aspect observable de la nature. Deuxièmement, classer ces observations d'une façon ordonnée (cela ne les modifie pas, mais les rend plus faciles à manipuler. De même, au bridge, le joueur qui range ses cartes par couleurs et par ordre de valeurs ne change pas leur valeur, n'en tire pas automatiquement le meilleur jeu possible ; mais il se place dans les meilleures conditions pour arriver à des coups logiques.) Troisièmement, tirer de ces observations classées un principe qui résume les observations.

Par exemple, nous pouvons observer que dans l'eau, le marbre coule à pic, le bois flotte, le fer coule, une plume flotte, le mercure coule, l'huile d'olive flotte, etc. Si nous faisons une liste de tous les objets qui coulent à pic, une autre de tous ceux qui flottent, et si nous cherchons une propriété qui distingue tous les objets d'une liste de tous ceux de l'autre, nous

conclurons : les objets plus denses que l'eau coulent, les objets moins denses que l'eau flottent.

A leur nouvelle manière d'étudier l'Univers, les Grecs donnèrent le nom de *philosophie* (littéralement « amour de la connaissance », c'est-à-dire « désir de savoir »...).

LA GÉOMÉTRIE ET LES MATHÉMATIQUES

C'est en géométrie que les Grecs ont obtenu leurs plus brillants succès. On peut attribuer ces succès à deux techniques principales : l'abstraction et la généralisation.

Voici un exemple. Les arpenteurs égyptiens avaient trouvé un moyen commode de tracer un angle droit : ils divisaient une corde en douze parties égales, et en faisaient un triangle dont les côtés valaient respectivement trois, quatre et cinq de ces parties. Entre le côté de trois et le côté de quatre, il y avait un angle droit. Nous ne savons pas comment les Égyptiens avaient découvert ce procédé ; apparemment, ils ne sont pas allés plus loin et se sont contentés de l'utiliser. Mais les Grecs eurent la curiosité de se demander pourquoi un tel triangle avait cette propriété. Au cours de leur analyse, ils se rendirent compte que la réalisation matérielle du triangle importait peu, qu'il pouvait être aussi bien en corde, en toile ou en lattes de bois. C'était seulement une propriété des « droites » formant un triangle. En concevant des droites idéales, qui soient indépendantes de toute réalisation matérielle, et existent seulement dans l'imagination, les Grecs ont inventé l'*abstraction* – technique consistant à éliminer tout ce qui n'est pas essentiel, et à considérer seulement les propriétés nécessaires à la solution du problème.

Les géomètres grecs réalisèrent un autre progrès en cherchant des solutions générales à des familles de problèmes, au lieu de traiter chaque problème séparément. Par exemple, on aurait pu découvrir par tâtonnement qu'il y avait aussi un angle droit dans les triangles de côtés 5,12 et 13, ainsi que dans les triangles de côtés 7, 24 et 25. Mais ce ne sont là que des nombres sans signification. Pouvait-on trouver une propriété commune à tous les triangles rectangles ? En raisonnant soigneusement, les Grecs montrèrent qu'un triangle est rectangle si et seulement si les longueurs de ses côtés satisfont à la relation $x^2 + y^2 = z^2$, z étant la longueur du plus grand côté. L'angle droit se trouve au point de rencontre des côtés de longueurs x et y. Ainsi, pour le triangle de côtés 3, 4 et 5, les carrés des côtés sont 9, 16 et 25 (et $9 + 16 = 25$). De même pour le triangle 5,12,13 ($25 + 144 = 169$) et pour le triangle 7,24 et 25 ($49 + 576 = 625$). Ce sont seulement trois cas particuliers parmi l'infinité de cas possibles, et en tant que tels, ils sont sans intérêt. Ce qui préoccupait les Grecs c'était de trouver la preuve que cette relation devait être vérifiée dans tous les cas. Et ils développèrent la géométrie comme manière élégante de découvrir et de formuler de telles généralisations.

Divers mathématiciens grecs proposèrent des démonstrations de propriétés des figures géométriques. Celle qui caractérise le triangle rectangle est attribuée à Pythagore, vers 525 avant notre ère, et porte encore son nom.

Vers 300 avant notre ère, Euclide rassembla les théorèmes mathématiques établis jusque-là, et les arrangea dans un ordre logique, de manière

que chaque théorème puisse être démontré à partir de ceux qui le précédaient. Bien sûr, en remontant cette chaîne de démonstration, on finit par arriver à quelque chose d'indémontrable : si chaque théorème se démontre à partir des précédents, comment démontrer le premier ? La solution est de commencer par postuler une vérité suffisamment évidente pour pouvoir se passer de preuve, ce que l'on appelle un *axiome*. Euclide parvint à réduire à quelques-uns les axiomes nécessaires à son époque, et à construire sur cette base le système compliqué et admirable de la *géométrie euclidienne*. On n'avait jamais tant construit à partir de si peu de choses, et à quelques modifications près les livres d'Euclide sont restés en usage pendant plus de deux mille ans.

LA DÉDUCTION

L'établissement d'un ensemble de connaissances comme conséquence inévitable d'un groupe d'axiomes (la *déduction*) est un jeu fascinant pour lequel les Grecs se passionnèrent, devant le succès de leur géométrie – au point de commettre deux sérieuses erreurs.

Tout d'abord, ils en vinrent à considérer la déduction comme le seul moyen respectable d'atteindre au savoir. Ils se rendaient pourtant compte qu'elle est inutilisable pour acquérir certains types de connaissances. Par exemple, la distance entre Athènes et Corinthe ne peut pas se déduire de principes abstraits ; il faut la mesurer. Les Grecs acceptaient d'observer la nature quand c'était nécessaire, mais ils en déploraient toujours la nécessité, et considéraient que le savoir le plus élevé était toujours atteint par le raisonnement. Ils tendaient à sous-estimer les connaissances liées à la vie de tous les jours. On raconte qu'un étudiant de Platon, que le maître instruisait en mathématiques, finit par demander avec impatience : « Mais à quoi sert tout cela ? » Platon, profondément offensé, aurait appelé un esclave, pour le charger de remettre une pièce de monnaie à l'étudiant, en disant : « Maintenant, vous ne penserez plus que cette instruction ne vous a rien apporté. » Après quoi l'étudiant fut mis à la porte.

On pense généralement que cette attitude supérieure venait du fait que la société grecque était fondée sur l'esclavage, et s'en remettait aux esclaves pour toutes les questions pratiques. C'est possible, mais j'ai plutôt tendance à penser que, pour les Grecs, la philosophie était un sport, un jeu intellectuel. Beaucoup de gens, en matière de sport, considèrent l'amateur comme un « gentleman » socialement supérieur au professionnel pour qui le sport est un métier. C'est au nom de ces idées puristes que nous prenons des précautions presque ridicules pour nous assurer que les concurrents des Jeux olympiques sont exempts de toute « tache » de professionnalisme. Les justifications rationnelles apportées par les Grecs à leur « culte de l'inutile » pourraient de la même façon avoir été fondées sur le sentiment qu'en permettant à des connaissances terre-à-terre (comme la distance d'Athènes à Corinthe) de venir se mêler à la pensée abstraite, on laissait l'imperfection pénétrer l'Eden de la vraie philosophie.

Mais quelles que soient les justifications qu'ils avançaient, les penseurs grecs furent sévèrement handicapés par cette attitude. La contribution grecque à la technique n'est pas nulle, mais même le plus grand des ingénieurs grecs, Archimède, a refusé d'écrire sur ses inventions et ses découvertes pratiques. Pour maintenir son statut d'amateur, il n'a diffusé

que ses travaux en mathématiques pures. Et le manque d'intérêt pour les choses terre-à-terre – qu'il s'agisse d'inventions, d'expériences ou d'observations – ne fut que l'un des facteurs qui enfermèrent la pensée grecque dans les limites étroites. La préférence des Grecs pour les études formelles et abstraites – en particulier, le succès de leur géométrie – les conduisit à une seconde erreur, et finalement à une impasse.

Séduits par la réussite des axiomes en géométrie, les Grecs en vinrent à considérer les axiomes comme des « vérités absolues », et à supposer que les autres branches du savoir pouvaient être élaborées à partir de vérités absolues du même genre. C'est ainsi qu'en astronomie, ils finirent par considérer comme des axiomes évidents par eux-mêmes les idées selon lesquelles la Terre était immobile au centre de l'Univers, et elle était corrompue et imparfaite, en opposition au ciel, considéré comme éternel, immuable et parfait. Puisque les Grecs considéraient le cercle comme la courbe parfaite, et que les cieux étaient parfaits, il s'ensuivait que tous les corps célestes devaient décrire des cercles autour de la Terre. Avec le temps, leurs observations (pour la navigation et l'établissement de calendriers) montrèrent que les planètes ne suivaient pas des trajectoires simplement circulaires, si bien que les Grecs durent représenter les mouvements des planètes par des combinaisons de plus en plus compliquées de cercles, aboutissant, vers l'an 150 de notre ère, au système complexe formulé à Alexandrie par Ptolémée. De la même façon, Aristote construisit des théories mécaniques fantaisistes, à partir d'axiomes « évidents par eux-mêmes », comme celui qui voulait que la vitesse de chute d'un objet soit proportionnelle à son poids (chacun savait qu'une pierre tombait plus vite qu'une plume).

Cet enthousiasme pour la déduction à partir d'axiomes devait aboutir à une impasse. Après que les Grecs eurent établi toutes les conséquences des axiomes, de nouvelles découvertes en mathématiques ou en astronomie semblaient hors de question. Le savoir philosophique semblait complet et parfait. Et pendant presque deux mille ans après l'âge d'or de la Grèce, quand une question se posait à propos du monde matériel, on avait tendance à satisfaire tout le monde en affirmant : « Aristote dit... » ou « Euclide dit... ».

LA RENAISSANCE ET COPERNIC

Ayant résolu les problèmes des mathématiques et de l'astronomie, les Grecs se tournèrent vers des domaines plus difficiles, en particulier l'âme humaine.

Platon s'intéressait beaucoup plus à des questions comme « Qu'est-ce que la justice ? » ou « Qu'est-ce que la vertu ? » qu'à savoir pourquoi la pluie tombe ou les planètes se déplacent. Étant le plus grand des Grecs pour la philosophie morale, il a supplanté Aristote, le plus grand pour la philosophie naturelle. Les penseurs grecs de l'époque romaine furent de plus en plus attirés par les délices de la philosophie morale, au détriment de la philosophie naturelle, apparemment frappée de stérilité. Le dernier développement de la philosophie antique fut le néo-platonisme, excessivement mystique, formulé par Plotin vers 250.

Le christianisme, insistant sur la nature de Dieu et sa relation avec l'homme, introduisit dans le domaine de la philosophie morale une dimension tout à fait nouvelle, augmentant encore son apparente supériorité, comme activité intellectuelle, sur la philosophie naturelle. De 200 à 1200, les Européens s'occupèrent à peu près exclusivement de philosophie morale, en particulier de théologie. La philosophie naturelle était pratiquement tombée dans l'oubli.

Mais les Arabes avaient réussi à préserver Aristote et Ptolémée tout au long du Moyen Age, et à travers eux la philosophie naturelle des Grecs finit par pénétrer à nouveau l'Europe occidentale. Vers 1200, on avait redécouvert Aristote. D'autres éléments arrivèrent par l'Empire byzantin agonisant, dernier bastion en Europe à maintenir une tradition culturelle continue depuis la grande époque de la Grèce.

La première, et la plus naturelle, des conséquences de la redécouverte d'Aristote fut l'application à la théologie de son système de raisonnement logique. Vers 1250, le théologien italien saint Thomas d'Aquin établit le système connu sous le nom de « thomisme », fondé sur des principes aristotéliciens, et qui représente encore la théologie fondamentale de l'Église catholique. Mais les Européens commencèrent bientôt à appliquer à des domaines plus concrets la renaissance de la pensée grecque.

Les chefs de file de la Renaissance mettant l'accent sur les réalisations humaines plutôt que sur les choses divines, on les appela « humanistes », et on donna le nom d'« humanités » aux études sur la littérature, l'art et l'histoire.

C'est d'un œil nouveau que les penseurs de la Renaissance retrouvèrent la philosophie naturelle des anciens Grecs, car les idées anciennes ne paraissaient plus tout à fait satisfaisantes. En 1543, l'astronome polonais Nicolas Copernic publia un livre qui allait jusqu'à rejeter l'axiome de base de l'astronomie : il proposait de placer au centre de l'Univers le Soleil, et non plus la Terre (mais il gardait l'idée de trajectoires circulaires pour la Terre et les autres planètes). Ce nouvel axiome permettait de décrire beaucoup plus simplement les mouvements observés des différents corps célestes. Cependant, l'axiome copernicien d'une Terre mobile était beaucoup moins « évident par lui-même » que l'axiome grec d'une Terre immobile, et il fallut plus d'un demi-siècle pour que la théorie de Copernic s'impose.

En un sens, le système copernicien n'était pas en lui-même un changement décisif. Copernic avait seulement remplacé un axiome par un autre. Et Aristarque de Samos avait déjà, deux mille ans plus tôt, proposé le même remplacement de la Terre par le Soleil au centre du monde. Non qu'un changement d'axiome soit sans importance. Quand les mathématiciens du XIX[e] siècle mirent en doute les axiomes d'Euclide et élaborèrent des géométries « non euclidiennes » fondées sur d'autres hypothèses, ils influencèrent profondément les idées sur de nombreux sujets ; aujourd'hui, l'histoire même et la forme de l'Univers semblent obéir plutôt à une géométrie non euclidienne qu'à la « géométrie du bon sens » qu'était celle d'Euclide. Mais la révolution déclenchée par Copernic n'entraînait pas seulement un changement d'axiomes ; elle impliquait une approche entièrement nouvelle des phénomènes naturels. Cette révolution, c'est l'Italien Galileo Galilei (dit Galilée) qui devait la mener à bien au tournant du XVI[e] siècle et au début du XVII[e].

EXPÉRIENCE ET INDUCTION

Les Grecs, dans l'ensemble, s'étaient contentés d'accepter les phénomènes « évidents » comme points de départ de leurs raisonnements. Il n'est rapporté nulle part qu'Aristote ait jamais laissé tomber deux pierres de poids différents pour vérifier son hypothèse que la vitesse de chute est proportionnelle au poids d'un objet. Pour les Grecs, l'idée de faire une expérience ne s'imposait pas. Elle risquait de gêner la déduction pure et d'en abîmer la beauté. De plus, si une expérience contredisait une déduction, comment être sûr que l'expérience était correcte ? Était-il vraisemblable que le monde imparfait de la réalité soit complètement en accord avec le monde parfait des idées abstraites ? Et s'il ne l'était pas, devait-on adapter le parfait aux exigences de l'imparfait ? La vérification d'une théorie parfaite par des instruments imparfaits ne semblait pas, pour les philosophes grecs, une manière valable d'acquérir des connaissances.

L'expérience a commencé à devenir philosophiquement respectable en Europe avec les travaux de philosophes comme Roger Bacon (contemporain de saint Thomas d'Aquin) et son homonyme plus récent Francis Bacon. Mais c'est Galilée qui renversa l'édifice grec et réalisa la révolution. C'était un logicien convaincant, et il avait le génie de la popularisation des idées. Il décrivit ses expériences et ses idées d'une façon si claire et si passionnante qu'il gagna à ses vues la communauté intellectuelle européenne. Et celle-ci accepta ses méthodes en même temps que ses résultats.

Suivant ce que l'on raconte, Galilée soumit la théorie d'Aristote sur la chute des corps à une expérience capable de faire du bruit dans toute l'Europe : il serait monté en haut de la tour penchée de Pise, et aurait lâché en même temps un boulet d'une livre et un boulet de dix livres. Le bruit des deux boulets frappant le sol exactement au même instant aurait tué la physique aristotélicienne...

En fait, Galilée n'a probablement pas fait cette expérience, mais elle est si typique de ses méthodes convaincantes qu'il n'est pas étonnant que cette histoire ait traversé les siècles.

En tout cas, Galilée a réellement fait rouler des boules le long de plans inclinés, en mesurant les distances qu'elles parcouraient pendant des intervalles de temps donnés. Il fut le premier à faire intervenir le temps dans ses expériences et à pratiquer des mesures systématiques.

Sa révolution a consisté à placer l'« induction » au-dessus de la déduction comme méthode scientifique. Au lieu de bâtir des conclusions à partir d'un ensemble d'hypothèses générales, la méthode inductive part des observations et en tire des généralisations (des axiomes, si l'on veut). Bien sûr, les Grecs aussi tiraient leurs axiomes de l'observation. L'axiome d'Euclide suivant lequel la ligne droite est le plus court chemin entre deux points était un énoncé intuitif fondé sur l'expérience. Mais tandis que le philosophe grec minimisait le rôle joué par l'induction, les scientifiques modernes y voient la manière essentielle de gagner des connaissances, le seul moyen de justifier les généralisations. De plus, le scientifique se rend compte qu'aucune généralisation n'est valide tant qu'elle n'a pas été vérifiée à maintes reprises par des expériences nouvelles – le test prolongé d'inductions nouvelles.

Le point de vue général actuel est juste l'opposé de celui des Grecs. Loin de considérer le monde réel comme une représentation imparfaite du vrai idéal, nous considérons les généralisations comme des représentations imparfaites du monde réel. Si ample et variée soit-elle, la vérification

inductive ne saurait rendre une généralisation complètement et absolument valide. Même si des milliards d'observations tendent à valider une généralisation, une seule observation contradictoire suffit pour que l'on soit obligé de la modifier. Et quel que soit le nombre de fois où une théorie est sortie victorieuse des vérifications expérimentales, on ne peut être sûr qu'elle survivra à la vérification suivante.

Voilà donc une pierre angulaire de la philosophie naturelle moderne. Celle-ci ne se propose pas d'atteindre une vérité ultime. En fait, l'expression « vérité ultime » est vide de sens, parce qu'il n'y a aucun moyen de faire assez d'observations pour rendre la vérité sûre, et donc « ultime ». Les philosophes grecs ne reconnaissaient aucune limite de ce genre. Qui plus est, ils ne voyaient aucune difficulté à appliquer exactement les mêmes méthodes de raisonnement aux questions « Qu'est-ce que la justice ? » et « Qu'est-ce que la matière ? ». Au contraire, la science moderne distingue nettement les deux types de questions. La méthode inductive ne saurait faire de généralisations à partir de ce qu'elle ne peut pas observer. Puisque la nature de l'âme humaine, par exemple, n'est directement observable par aucun moyen actuellement connu, elle échappe au domaine d'action de la méthode inductive.

La victoire de la science moderne n'est devenue complète qu'après avoir établi un autre principe essentiel, la liberté de communication et de coopération entre tous les scientifiques. Bien que cette nécessité nous semble maintenant évidente, elle ne l'était pas pour les philosophes de l'Antiquité et du Moyen Age. Les Pythagoriciens de la Grèce antique formaient une société secrète, qui gardait pour elle ses découvertes mathématiques. Les alchimistes du Moyen Age rendaient délibérément leurs écrits incompréhensibles, pour garder dans un cercle aussi restreint que possible leurs soi-disant découvertes. Au XVI[e] siècle, le mathématicien italien Niccolo Tartaglia, ayant découvert une méthode de résolution des équations du troisième degré, trouva naturel d'essayer d'en garder le secret. Quand Gerolamo Cardano, l'un de ses collègues mathématiciens, après lui avoir soutiré son secret en promettant de le garder, s'empressa de le publier, Tartaglia en éprouva une fureur assez naturelle. Mais si on ferme les yeux sur la duplicité de Cardano, il avait parfaitement raison d'estimer qu'il fallait publier une telle découverte.

De nos jours, aucune découverte scientifique n'est considérée comme telle si elle est gardée secrète. Le chimiste anglais Robert Boyle, un siècle après Tartaglia et Cardano, a souligné l'importance de la publication détaillée de tous les résultats scientifiques. Qui plus est, on n'admet la validité d'une observation, même publiée, qu'après sa vérification par au moins un autre chercheur. La science n'est pas produite par des individus, mais par la « communauté scientifique ».

L'un des premiers groupes (et certainement le plus célèbre) à représenter une telle communauté scientifique fut la Royal Society (dont le nom complet était « Société royale de Londres pour l'amélioration de la connaissance de la nature »). Elle naquit des rencontres informelles, à partir de 1645, d'un groupe de « gentlemen » intéressés par les nouvelles méthodes scientifiques introduites par Galilée. En 1660, la société reçut son statut officiel du roi Charles II.

Les membres de la Royal Society se rencontraient, discutaient ouvertement les résultats de leurs travaux, écrivaient des lettres pour les décrire, en anglais plutôt qu'en latin, et poursuivaient vigoureusement

leurs expériences. Pourtant, pendant la plus grande partie du XVIIe siècle, ils restèrent dans une attitude défensive. On pourrait représenter l'attitude de la plupart de leurs collègues contemporains par un dessin humoristique à la manière moderne, où l'on verrait les ombres majestueuses de Pythagore, Euclide et Aristote regarder de haut des gamins en train de jouer aux billes et portant l'étiquette « Royal Society ».

Tout devait changer avec les travaux d'Isaac Newton, qui devint membre de la société. A partir des observations et des conclusions de Galilée, ainsi que des observations astronomiques du Danois Tycho Brahé, dont l'astronome allemand Johannes Kepler avait tiré l'idée des orbites elliptiques pour les planètes, Newton arriva par induction à ses trois lois simples du mouvement et à sa grande généralisation fondamentale : la loi de la gravitation universelle. (Pourtant, quand il publia ses résultats, il utilisa la géométrie et la méthode grecque d'explication déductive.) Cette découverte impressionna tellement le monde savant que Newton fut idolâtré, presque déifié de son vivant. Devant ce nouvel univers majestueux, fondé sur quelques lois simples tirées de raisonnements inductifs, c'étaient maintenant les philosophes grecs qui avaient l'air d'un groupe de gamins en train de jouer aux billes. La révolution, déclenchée par Galilée au début du XVIIe siècle, triomphait avec Newton à la fin du même siècle.

LA SCIENCE MODERNE

On aimerait pouvoir dire que la science et les hommes ont « vécu heureux » ensemble depuis cette époque. Mais en fait, les vraies difficultés ne faisaient que commencer. Tant que la science était restée déductive, la philosophie naturelle pouvait faire partie de la culture générale de tous les hommes instruits (les femmes, hélas, n'ayant accédé à l'instruction, sauf exceptions, que très récemment). Mais la science inductive devint un travail considérable d'observation, d'étude et d'analyse. Ce n'était plus un divertissement pour les amateurs. Et la science devenait plus complexe d'année en année. Pendant un siècle après Newton, il était encore possible pour un homme exceptionnel de maîtriser tous les domaines de connaissances scientifiques. Mais dès 1800, c'était devenu tout à fait impossible. Et il était de plus en plus nécessaire, pour un scientifique, de se limiter au domaine étroit qu'il approfondissait. C'est sa croissance même qui a inexorablement contraint la science à se spécialiser. Et chaque génération de scientifiques était plus spécialisée que la précédente.

Les publications scientifiques de travaux personnels n'ont jamais été aussi nombreuses, ni aussi illisibles pour quiconque, sauf pour un spécialiste de la même branche. Ceci a constitué un grand handicap pour la science elle-même, car les progrès fondamentaux de la connaissance scientifique résultent souvent du « mariage » de connaissances issues de branches différentes. Encore plus grave : la science perd de plus en plus contact avec les non-scientifiques. Dans ces conditions, on en vient à considérer les scientifiques presque comme des sorciers, inspirant la crainte plutôt que l'admiration. Et le sentiment que la science est une magie incompréhensible, sauf par une poignée d'élus présentant avec le reste de l'humanité des différences inquiétantes, ne peut que détourner beaucoup de jeunes de la science.

Depuis la Seconde Guerre mondiale, on trouve parmi les jeunes – même ceux qui fréquentent les universités – une forte et franche hostilité à l'égard de la science. Notre société industrialisée est fondée sur les découvertes scientifiques des deux cents dernières années, et se trouve confrontée à des fléaux qui sont le revers de la médaille technologique.

L'amélioration des techniques médicales a entraîné une prolifération galopante de la population. Les industries chimiques et les moteurs à explosion polluent notre air et notre eau. La consommation de matières premières et d'énergie épuise et détruit l'écorce terrestre. Tous ces maux sont trop facilement imputés à la « science » et aux « scientifiques » par ceux qui ne comprennent pas bien que si le savoir peut créer des problèmes, ce n'est pas l'ignorance qui les résoudra.

Pourtant, il n'est pas inévitable que la science paraisse aussi mystérieuse aux non-scientifiques. Un grand pas serait fait pour combler la barrière si les scientifiques acceptaient la responsabilité de communiquer – d'expliquer leur domaine aussi simplement que possible au plus grand nombre possible de gens – et si les non-scientifiques, pour leur part, acceptaient leur responsabilité : écouter. Pour se faire une idée satisfaisante des progrès dans un domaine scientifique donné, il n'est pas du tout indispensable de comprendre complètement la science. Après tout, personne ne pense qu'il faille être capable d'écrire une grande œuvre littéraire pour apprécier Shakespeare, et pour écouter avec plaisir une symphonie de Beethoven, l'auditeur n'est pas obligé de savoir en composer une équivalente. De la même façon, on peut apprécier les progrès scientifiques et y prendre plaisir, même si on n'a personnellement aucune disposition pour la créativité scientifique.

Mais, demandera-t-on, à quoi cela servirait-il ? La première réponse est que personne ne peut se sentir chez soi dans le monde moderne, juger la nature de ses problèmes et les solutions possibles, à moins d'avoir une idée intelligente de ce que nous réserve la science. De plus, l'initiation au monde magnifique de la science apporte une intense satisfaction esthétique, inspire la jeunesse, comble le désir de savoir et permet d'apprécier plus profondément les merveilles déjà réalisées par l'esprit humain, et celles dont il est capable.

C'est pour permettre une telle initiative que j'ai entrepris d'écrire ce livre.

PREMIÈRE PARTIE

Les sciences physiques

Chapitre 2

L'Univers

La taille de l'Univers

Si on jette un coup d'œil vers le ciel, rien ne le fait paraître particulièrement lointain. Les jeunes enfants acceptent sans problème les fables où « la vache saute par-dessus la Lune », ou l'idée que quelqu'un puisse « sauter assez haut pour toucher le ciel ». Les Grecs anciens, à l'époque des mythes, ne voyaient rien de ridicule à faire reposer le ciel sur les épaules du géant Atlas. On pourrait, bien sûr, attribuer à Atlas une taille prodigieuse, d'ordre astronomique, mais un autre mythe nous en empêche. Pour le onzième de ses « douze travaux », aller chercher les pommes d'or (oranges ?) des Hespérides (loin vers l'ouest : l'Espagne, peut-être ?), Hercule embaucha Atlas. Et pendant que celui-ci allait chercher les pommes, Hercule, debout sur le sommet d'une montagne, soutint le ciel à sa place. Or, si Hercule était à coup sûr robuste, ce n'était pourtant pas un géant. Par conséquent, les premiers Grecs admettaient sans peine que le ciel ne soit qu'à quelques mètres du sommet des montagnes.

Il est naturel, pour commencer, de considérer le ciel comme une voûte solide, où les corps célestes brillants sont enchâssés comme des diamants (c'est ainsi que la Bible l'appelle « firmament », d'après un mot latin qui a donné « ferme », solide). Mais, dès la période qui va du VIe au IVe siècle av. J.-C., les astronomes grecs réalisèrent qu'il devait y avoir plusieurs voûtes. En effet, tandis que les « étoiles fixes » tournaient en bloc autour de la Terre, apparemment sans changer leurs positions relatives, il n'en

allait pas de même du Soleil, de la Lune, et de cinq « étoiles » brillantes : Mercure, Vénus, Mars, Jupiter et Saturne. En fait, chacun de ces corps se déplaçait sur son chemin particulier. Aussi ces sept corps reçurent-ils le nom de *planètes* (« errantes » en grec), et il semblait évident qu'ils ne pouvaient pas être fixés à la voûte qui portait les étoiles.

Pour les Grecs, chaque planète était enchâssée dans sa propre voûte sphérique invisible, les voûtes étant placées les unes à l'intérieur des autres, la plus proche appartenant à la planète qui se déplaçait le plus vite. C'était la Lune, qui faisait le tour du ciel en vingt-sept jours un tiers à peu près. Ensuite venaient dans l'ordre (d'après les Grecs) : Mercure, Vénus, notre Soleil, Mars, Jupiter et Saturne.

LES PREMIÈRES MESURES

La première mesure scientifique d'une dimension à l'échelle du monde date d'environ 240 ans av. J.-C. Ératosthène, directeur de la bibliothèque d'Alexandrie, alors la plus puissante des institutions scientifiques du monde, remarqua que le 21 juin, à midi, le Soleil était exactement à la verticale à Syène, en Égypte, tandis qu'à Alexandrie, située à huit cents kilomètres plus au nord, il n'était pas tout à fait au zénith au même moment. Ératosthène se servit de cela pour calculer le rayon de la Terre. A partir de la longueur des ombres à midi à Alexandrie, le jour du solstice, et en admettant que la Terre était sphérique – chose déjà admise par les astronomes grecs de l'époque – un calcul géométrique permettait de trouver la circonférence de la Terre, et par conséquent son diamètre (figure 2.1).

Le calcul d'Ératosthène, exprimé en unités grecques, donnait environ une circonférence de 40 000 kilomètres, ce qui est d'une justesse étonnante. Malheureusement, cette valeur correcte ne resta pas en vigueur : une centaine d'années avant notre ère, un autre astronome grec, Posidonius, recommença le travail d'Ératosthène et trouva seulement, pour la circonférence de la Terre, une valeur de 30 000 kilomètres.

C'est cette valeur qui fit autorité tout au long de l'Antiquité et du Moyen Age. C'est à partir de cette valeur que Christophe Colomb estima à 5 000 kilomètres la distance qui devait le mener en Asie. S'il avait connu la vraie taille de la Terre, peut-être aurait-il renoncé à son voyage. C'est seulement en 1521-1523, quand la flotte de Magellan (ou plutôt le seul navire survivant de cette flotte) fit le tour du monde, que l'on adopta finalement pour la circonférence de la Terre la valeur correcte donnée par Ératosthène.

Environ en 150 av. J.-C., Hipparque de Nicée calcula, en fonction du diamètre de la Terre, la distance qui nous sépare de la Lune. Il utilisa une méthode suggérée cent ans plus tôt par Aristarque de Samos, le plus hardi des astronomes grecs. Les Grecs savaient déjà que les éclipses de Lune étaient causées par l'ombre de la Terre, quand celle-ci passe exactement entre le Soleil et la Lune. Aristarque se rendit compte que la courbure du bord de cette ombre projetée sur la surface de la Lune permettait de comparer la taille de la Lune à celle de la Terre. A partir de là, des méthodes géométriques permettaient de calculer la distance Terre-Lune en fonction du diamètre de la Terre. Reprenant ce travail, Hipparque calcula que la distance Terre-Lune valait trente fois

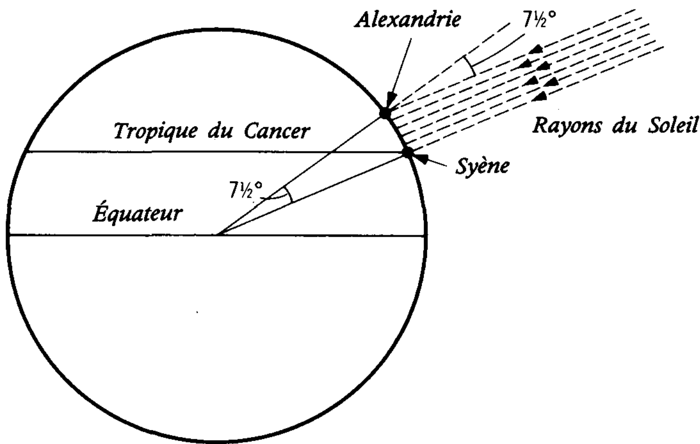


Figure 2.1. Ératosthène a mesuré la taille de la Terre à partir de sa courbure. A midi, le 21 juin, les rayons du Soleil sont verticaux à Syène, qui se trouve sur le tropique du Cancer. Mais au même moment, à Alexandrie, qui est plus au nord, ces rayons font un angle de 7,5 degrés avec la verticale, et projettent des ombres. Connaissant la longueur de ces ombres et la distance entre les deux villes, Ératosthène a pu faire son calcul.

le diamètre de la Terre. Si on garde pour celui-ci la valeur de 12 000 kilomètres trouvée par Ératosthène, on obtient donc pour la distance Terre-Lune la valeur de 360 000 kilomètres, qui est très proche de la valeur correcte.

Mais cette mesure de la distance Terre-Lune fut le dernier succès des astronomes grecs dans l'évaluation de la taille de l'Univers. Pourtant, Aristarque avait fait une tentative héroïque pour déterminer la distance Terre-Soleil. La méthode géométrique qu'il avait choisie était tout à fait correcte en théorie, mais elle exigeait la mesure d'angles si petits que, sans instruments modernes, elle ne permettait pas d'obtenir une valeur correcte. Pour Aristarque, le Soleil était à peu près 20 fois plus loin que la Lune (alors qu'il est en réalité 400 fois plus loin). Mais si mauvais que fussent ces résultats, Aristarque en avait déduit que le Soleil devait être au moins sept fois plus gros que la Terre. Trouvant illogique de voir un gros corps tourner autour d'un petit, il avait décidé que la Terre devait tourner autour du Soleil.

Malheureusement, personne ne l'écouta. Les astronomes qui lui succédèrent, d'Hipparque à Ptolémée, calculèrent tous les mouvements célestes sur la base d'une Terre immobile au centre de l'Univers, avec la Lune à 360 000 kilomètres, et les autres objets plus loin, à des distances inconnues. Ce modèle devait durer jusqu'en 1543, c'est-à-dire jusqu'à la publication du livre de Nicolas Copernic, retournant au point de vue d'Aristarque, et détrônant définitivement la Terre de sa place au centre de l'Univers.

LA MESURE DU SYSTÈME SOLAIRE

Le fait de placer le Soleil au centre du système ne permettait pas pour autant d'en déterminer les dimensions. Copernic adopta la valeur des Grecs pour la distance Terre-Lune, mais il n'avait aucune idée de l'éloignement du Soleil. C'est seulement en 1650 qu'un astronome belge, Godefroy Wendelin, recommença les observations d'Aristarque avec de meilleurs instruments, et trouva qu'au lieu d'être 20 fois plus loin que la Lune, il était 240 fois plus loin, c'est-à-dire environ à cent millions de kilomètres. C'était une valeur trop faible encore, mais bien meilleure que la précédente.

Entre-temps, en 1609, l'astronome allemand Johannes Kepler ouvrait la voie à des mesures plus précises, avec sa découverte que les orbites des planètes étaient des ellipses et non des cercles. Pour la première fois, il devenait possible de calculer correctement ces orbites, et de plus de représenter le système solaire à l'échelle, c'est-à-dire d'en calculer exactement les proportions ; dès lors, il suffisait de mesurer une seule distance entre deux planètes pour connaître aussitôt toutes les autres. Par conséquent, on n'avait plus besoin de mesurer directement la distance Terre-Soleil, comme avaient essayé de le faire Aristarque et Wendelin. Il suffisait de mesurer la distance entre la Terre et n'importe quel corps plus proche, autre que la Lune, par exemple Mars ou Vénus.

Une des méthodes possibles pour mesurer de telles distances repose sur l'utilisation de la *parallaxe.* Un exemple simple permet de comprendre de quoi il s'agit. Placez un doigt à dix centimètres devant vos yeux, et regardez-le tantôt avec l'œil gauche, tantôt avec le droit. Le doigt apparaît à des endroits différents par rapport au fond, parce que vous déplacez votre point de vue. Si vous recommencez en tenant le doigt plus loin, à bout de bras par exemple, le doigt se déplace encore par rapport au fond, mais beaucoup moins. On peut se servir de l'importance de ce déplacement pour évaluer la distance entre les yeux et le doigt.

Bien sûr, pour un objet situé à vingt mètres, le déplacement par rapport au fond est trop petit pour être mesurable ; il nous faut une « base » plus large que la distance entre nos deux yeux. Mais pour élargir cette base, il suffit de viser l'objet depuis un endroit donné, puis de se déplacer de cinq mètres vers la droite et de recommencer. La parallaxe est alors assez grande pour être mesurée, et on peut déterminer la distance à laquelle se trouve l'objet. C'est la méthode utilisée par les arpenteurs pour évaluer des distances à travers une rivière ou un ravin.

On peut employer la même méthode pour évaluer la distance de la Lune, en se servant des étoiles comme fond. Vue depuis un observatoire en Californie, par exemple, la Lune occupe une certaine position parmi les étoiles. Vue au même instant d'un observatoire en Angleterre, elle occupe une position légèrement différente. A partir de cette différence, et de la distance qui sépare les deux observatoires (en ligne droite, à travers la Terre), on peut calculer la distance de la Lune. Bien sûr, on peut en théorie élargir la base en faisant les observations à partir d'observatoires situés aux antipodes l'un de l'autre. Leur distance vaut alors le diamètre terrestre (près de 13 000 kilomètres), et l'angle de parallaxe obtenu, divisé par deux, donne la *parallaxe géocentrique.*

Le décalage entre les deux positions d'un corps céleste se mesure en degrés, minutes et secondes. Un degré vaut 1/360 de tour, chaque degré

est divisé en 60 minutes, et chaque minute en 60 secondes. Une minute d'arc vaut donc 1/21 600 tour du ciel, et une seconde soixante fois moins, soit 1/1 296 000 tour.

En utilisant la trigonométrie (les relations entre les côtés et les angles des triangles), Ptolémée était déjà parvenu à mesurer la distance de la Lune à partir de sa parallaxe, retrouvant ainsi la valeur calculée plus tôt par Hipparque. En fait, la parallaxe géocentrique de la Lune vaut 57 minutes (presque un degré). Le décalage est le même que celui d'un objet situé à trois mètres, et vu tantôt de l'œil droit, tantôt de l'œil gauche. Mais quand il s'agit de mesurer la parallaxe du Soleil ou d'une planète, les angles deviennent trop petits, et les astronomes grecs durent donc se contenter de signaler que ces corps étaient plus éloignés que la Lune – de combien, personne ne pouvait le savoir.

La trigonométrie seule, malgré son perfectionnement par les Arabes durant le Moyen Age et par les mathématiciens européens au XVI[e] siècle, ne suffisait pas à donner la réponse. Mais la mesure d'angles très petits devint possible avec l'invention de la lunette (c'est en 1609 que Galilée construisit la première lunette sérieuse et la tourna vers le ciel, après avoir entendu parler de l'instrument inventé quelques mois plus tôt par un lunetier hollandais).

La méthode des parallaxes atteignit enfin un objet plus éloigné que la Lune, en 1673, quand l'astronome français d'origine italienne Jean-Dominique Cassini détermina la parallaxe de Mars. Il repéra la position de Mars par rapport aux étoiles en même temps que l'astronome français Jean Richer, en Guyane française. La combinaison des deux observations donna à Cassini la parallaxe de Mars et lui permit donc de calculer la taille du système solaire : il obtint, pour la distance Soleil-Terre, une valeur de 140 millions de kilomètres, à peine 7 % de moins que la valeur correcte.

Depuis lors, diverses parallaxes d'objets du système solaire ont été mesurées, avec des précisions croissantes. En 1931, un vaste programme international visa à déterminer la parallaxe de la petite planète Éros, qui approcha la Terre, cette année-là, de plus près qu'aucun autre objet sauf la Lune. Aussi sa parallaxe était grande, ce qui permit de la mesurer avec une grande précision et de déterminer ainsi, plus exactement qu'auparavant, la taille du système solaire. A partir de ces calculs, et en utilisant des méthodes encore plus précises que celles qui sont fondées sur la parallaxe, on attribue maintenant à la distance Soleil-Terre une valeur moyenne de 149 millions de kilomètres (l'orbite de la Terre étant elliptique, cette distance varie légèrement de part et d'autre de cette valeur moyenne).

On donne à cette distance moyenne le nom d'*unité astronomique* (UA), et on s'en sert pour exprimer les autres distances à l'intérieur du système solaire. Saturne, par exemple, est à une distance moyenne du Soleil de 1,4 milliard de kilomètres, soit 9,54 UA. Au fur et à mesure des découvertes des planètes extérieures – Uranus, Neptune et Pluton – les frontières du système solaire s'éloignaient. Le diamètre maximum de l'orbite de Pluton vaut 11,8 milliards de kilomètres, soit 79 UA. Et certaines comètes s'éloignent encore plus du Soleil.

Dès 1830, on savait que la taille du système s'évalue en milliards de kilomètres. Mais de toute évidence cela n'est rien devant la taille de l'Univers : il reste les étoiles...

PLUS LOIN, LES ÉTOILES

On peut, bien sûr, continuer à imaginer les étoiles comme de petits objets enchâssés dans la voûte solide du ciel, celle-ci constituant, juste autour du système solaire, la limite de l'Univers. Jusqu'au début du XVIIIe siècle, cette idée resta tout à fait respectable, en dépit de quelques savants qui commençaient à la mettre en doute.

Dès 1440, l'Allemand Nicolas de Cuse soutint que l'espace était infini, et que les étoiles étaient des soleils, qui s'étendaient dans toutes les directions sans limite, chacune avec son cortège de planètes habitées. C'est leur grand éloignement qui les fait apparaître non pas comme des soleils, mais comme des points lumineux. Malheureusement, ces idées ne s'appuyaient sur aucune preuve, et Nicolas de Cuse les avança seulement comme des hypothèses. Elles paraissaient délirantes, et on les ignora.

Mais en 1718, l'astronome anglais Edmund Halley s'appliqua à déterminer avec précision la position des étoiles, et constata que trois des plus brillantes – Sirius, Procyon et Arcturus – n'étaient plus à l'emplacement repéré par les astronomes grecs. Même en tenant compte du fait que ceux-ci observaient à l'œil nu, la différence était trop grande pour résulter d'une erreur d'observation. Halley en déduisit que les étoiles, après tout, n'étaient pas fixées rigidement au firmament, mais qu'elles se déplaçaient indépendamment les unes des autres, comme des abeilles dans un essaim. Ce mouvement est très lent, et si peu évident qu'on ne put pas le déceler avant l'invention des lunettes.

Ce qui rend si petit ce *mouvement propre,* c'est la distance considérable à laquelle sont situées les étoiles. Sirius, Procyon et Arcturus se trouvent parmi les plus proches, et c'est pourquoi leur mouvement propre a fini par devenir appréciable. C'est aussi leur proximité qui les rend si brillantes. Les étoiles qui nous semblent moins lumineuses sont en général plus lointaines, et leur mouvement propre reste imperceptible, même au bout d'un temps aussi long que celui qui nous sépare des anciens Grecs.

S'il révèle bien les distances variables des étoiles, le mouvement propre ne permet pas pour autant de calculer ces distances. Bien sûr, les étoiles les plus proches devraient avoir une parallaxe par rapport aux plus éloignées. Mais on n'a jamais pu détecter cette parallaxe. Même en utilisant comme base le diamètre de l'orbite terrestre (300 millions de kilomètres), les astronomes ne sont pas parvenus à observer une parallaxe d'étoile. Ceci veut dire que même les étoiles les plus proches se trouvent à des distances considérables. Et à mesure que les télescopes se perfectionnaient, sans toutefois permettre d'apprécier cette parallaxe, on augmentait l'estimation minimale de ces distances. Pour être visibles à des distances aussi énormes, il fallait que les étoiles soient de gigantesques boules de feu, comme notre Soleil ; Nicolas de Cuse avait raison.

Mais les instruments continuaient à s'améliorer. Vers 1830, l'astronome allemand Friedrich Wilhelm Bessel utilisa un appareil d'invention récente, *l'héliomètre,* destiné initialement à mesurer avec une grande précision le diamètre du Soleil. On pouvait aussi s'en servir pour mesurer la distance entre deux étoiles. En mesurant au fil des mois les variations de ces distances, Bessel réussit finalement à mesurer une parallaxe d'étoile (figure 2.2). Il choisit une petite étoile de la constellation du Cygne, appelée 61 Cygni, parce qu'elle avait un mouvement propre remarquablement important par rapport aux autres étoiles, et qu'elle devait donc être

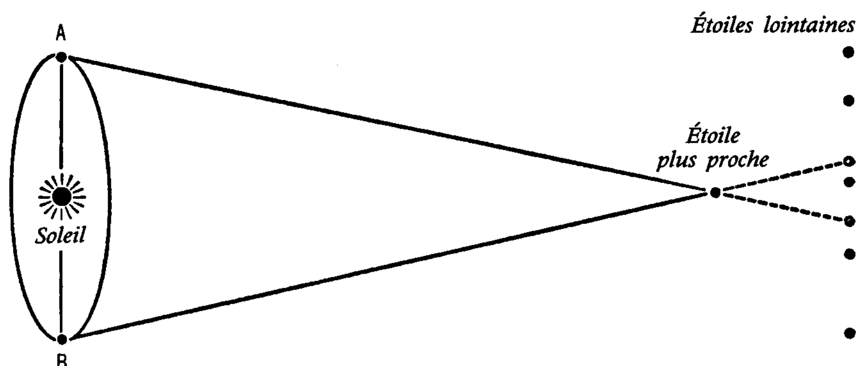


Figure 2.2. Parallaxe d'une étoile, mesurée à partir de points opposés sur l'orbite de la Terre autour du Soleil.

relativement proche de nous. (Cette dérive constante est le véritable mouvement propre, qu'il ne faut pas confondre avec les oscillations apparentes dues à l'effet de parallaxe.) Bessel repéra donc les positions successives de 61 Cygni par rapport au fond constitué par les étoiles voisines – vraisemblablement plus distantes – et poursuivit ses observations pendant plus d'un an. Enfin, en 1838, il annonça que 61 Cygni avait une parallaxe de 0,31 seconde d'arc – la largeur d'une pièce de monnaie à une vingtaine de kilomètres ! Cette parallaxe, mesurée avec le diamètre de l'orbite terrestre comme base, signifiait que 61 Cygni était à environ 100 000 milliards de kilomètres de nous – environ dix mille fois les dimensions du système solaire. Celui-ci donc, comparé aux distances des étoiles, même les plus proches, devenait absolument minuscule.

De telles distances, exprimées en milliards de kilomètres, ne sont pas commodes. Les astronomes préfèrent les exprimer en utilisant une autre unité de distance, l'*année-lumière*. C'est la distance parcourue en un an par la lumière dans le vide, à raison de 300 000 km/s. Exprimée dans cette nouvelle unité, la distance de 61 Cygni devient environ 11 années-lumière.

Deux mois après la victoire de Bessel, et perdant ainsi de peu l'honneur d'être le premier, l'astronome anglais Thomas Henderson annonça la distance de l'étoile Alpha Centauri. Située dans l'hémisphère sud, et invisible au nord de la latitude de Tampa (Floride), cette étoile est la troisième de notre ciel, par ordre de luminosité. On constata qu'elle avait une parallaxe de 0,75 seconde d'arc, plus de deux fois supérieure à celle de 61 Cygni. Elle était donc plus de deux fois plus proche : située seulement à 4,3 années-lumière du Soleil, c'est notre voisine la plus proche. En fait, ce n'est pas une seule étoile, mais un groupe de trois.

En 1840, un astronome russe d'origine allemande, Friedrich Wilhelm von Struve, annonça la parallaxe de Véga, la quatrième étoile brillante de notre ciel. On s'apercevra plus tard que sa mesure était assez inexacte, ce qui se comprend facilement : la parallaxe de Véga est très petite, l'étoile se trouvant à 27 années-lumière du Soleil.

A la fin du XIXe siècle, on avait ainsi déterminé, par la méthode des parallaxes, les distances de soixante-dix étoiles (et de plusieurs milliers en

1980). Même avec les meilleurs instruments, on ne peut pas dépasser une distance de cent années-lumière si on souhaite la moindre précision. Au-delà, il reste d'innombrables étoiles plus éloignées.

A l'œil nu, on peut compter environ 6 000 étoiles. Mais dès l'invention de la lunette, il devint évident que ce n'était là qu'une partie insignifiante des étoiles de l'Univers. Quand Galilée tourna sa lunette vers le ciel, en 1609, non seulement il découvrit des étoiles jusque-là invisibles, mais en pointant son instrument vers la Voie lactée, il éprouva un choc encore plus grand. A l'œil nu, la Voie lactée apparaît comme une bande de brouillard lumineux. A travers la lunette de Galilée, cette bande de brouillard se résolvait en milliards d'étoiles.

Cette multitude resta incompréhensible jusqu'en 1785 : cette année-là, un astronome anglais d'origine allemande, William Herschel, imagina que les étoiles que nous voyons étaient groupées pour former une sorte de lentille. En effet, nous voyons beaucoup d'étoiles si nous nous tournons dans la direction de la Voie lactée, et peu dans une direction à angle droit avec cette « roue » d'étoiles. Herschel en déduisit que les étoiles formaient un ensemble aplati, le plan d'aplatissement étant celui de la Voie lactée. Nous savons maintenant que cette idée est, à peu de choses près, correcte, et nous appelons notre groupe d'étoiles la *Galaxie*, un autre mot pour Voie lactée (« galaxie » vient du mot grec signifiant « le lait »).

Herschel tenta d'estimer la taille de la Galaxie. Il supposa que toutes les étoiles avaient à peu près la même luminosité intrinsèque, ce qui permettait d'évaluer la distance d'une étoile à partir de sa luminosité apparente. En effet, il est bien connu que celle-ci diminue en fonction du carré de la distance : si toutes les étoiles ont la même luminosité intrinsèque, et si l'étoile A nous paraît neuf fois moins lumineuse que B, elle est trois fois plus éloignée.

En comptant les étoiles dans plusieurs petits secteurs de la Voie lactée, Herschel estima à cent millions le nombre total des étoiles de la Galaxie. A partir des luminosités apparentes, il décida que le diamètre de la Galaxie valait 850 fois la distance qui nous sépare de l'étoile brillante Sirius, et que l'épaisseur de la Galaxie valait 155 fois cette distance.

Nous savons maintenant que Sirius est à 8,8 années-lumière, ce qui donne à la Galaxie de Herschel un diamètre de 7 500 années-lumière et une épaisseur de 1 300. En fait, ces estimations étaient bien inférieures aux valeurs réelles, mais, comme dans la mesure trop modeste de la distance du Soleil faite par Aristarque, c'était un pas dans la bonne direction.

Il était facile d'imaginer que dans la Galaxie les étoiles se déplaçaient d'une façon désordonnée, telles des abeilles dans un essaim (comme je l'ai déjà dit). Herschel montra que le Soleil lui-même se déplaçait de cette façon.

En 1805, après avoir passé vingt ans à déterminer les mouvements d'étoiles aussi nombreuses que possible, il constata que, dans une région du ciel, les étoiles semblaient en général s'écarter d'un point central (l'*apex*), tandis que dans une autre, située à l'opposé de la première, elles semblaient au contraire converger vers un autre centre (l'*anti-apex*).

L'explication la plus simple du phénomène consistait à supposer que le Soleil se déplaçait de l'anti-apex à l'apex, les étoiles semblant s'écarter devant à mesure qu'il s'en approchait, et se rapprocher derrière à mesure qu'il s'en éloignait. (Il se passe la même chose avec les arbres quand nous

traversons une forêt à pied, mais cela nous paraît si naturel que nous ne le remarquons même pas.)

Ainsi, le Soleil n'est pas, comme l'avait pensé Copernic, le centre immobile de l'Univers : il se déplace. Mais ce mouvement ne l'entraîne pas autour de la Terre, comme le croyaient les Anciens ; il emmène avec lui la Terre et toutes les planètes dans son voyage à travers la Galaxie. Les mesures modernes montrent que le Soleil se dirige vers un point dans la constellation de la Lyre, à une vitesse de 20 km/s.

A partir de 1906, l'astronome hollandais Jacobus Cornelis Kapteyn commença un nouvel « arpentage » du ciel. Disposant de la photographie et connaissant les distances réelles des étoiles les plus proches, il parvint à une estimation plus correcte que celle de Herschel : selon Kapteyn, la Galaxie avait 23 000 années-lumière de diamètre, et 6 000 d'épaisseur. Ainsi, la Galaxie de Kapteyn était quatre fois plus large et cinq fois plus épaisse que celle de Herschel. Pourtant, ces estimations étaient encore trop faibles.

En résumé, au début du XX[e] siècle on s'est trouvé, pour la mesure des distances des étoiles, dans la situation où l'on se trouvait en 1700 pour la mesure des distances dans le système solaire. En 1700, on connaissait la distance de la Lune, mais on était réduit à deviner celle des planètes. En 1900, on connaissait les distances des étoiles les plus proches, mais on était réduit à deviner celles des étoiles plus éloignées.

MESURER LA LUMINOSITÉ D'UNE ÉTOILE

Ce qui allait permettre d'aller plus loin, ce fut la découverte d'un nouvel « instrument » de mesure : certaines étoiles ont une luminosité qui change avec le temps. Ce nouveau chapitre commence avec une étoile assez brillante, appelée Delta Cephei, dans la constellation de Céphée. Étudiée avec soin, cette étoile révéla un cycle de luminosité variable : à partir de son éclat le plus faible, elle devient rapidement deux fois plus lumineuse, puis recommence à décliner lentement jusqu'au point de départ. Ces variations se reproduisent indéfiniment avec une régularité parfaite. Les astronomes découvrirent un grand nombre d'autres étoiles qui variaient de cette façon régulière, et en l'honneur de Delta Cephei on les appela *variables céphéides* ou simplement *céphéides.*

La période des céphéides (le temps entre deux minima successifs) varie d'un peu moins d'un jour à près de deux mois. Les plus proches de notre Soleil semblent avoir une période voisine d'une semaine. Celle de Delta Cephei vaut 5,3 jours, tandis que la plus proche de toutes les céphéides (l'Étoile polaire) a une période de 4 jours (mais sa variation relative de luminosité est trop faible pour être décelable à l'œil nu).

Pour apprécier l'importance des céphéides pour les astronomes, une petite digression s'impose au sujet de la luminosité.

C'est Hipparque qui a inventé, pour repérer la luminosité des étoiles, une échelle encore en usage : celle des *magnitudes*. Plus l'étoile est brillante, plus sa magnitude est petite. Hipparque attribua la *première magnitude* (on trouve aussi le terme de « première grandeur ») aux vingt étoiles les plus brillantes du ciel. Ensuite, la magnitude augmente à mesure que l'éclat diminue, jusqu'à la *sixième magnitude,* qui correspond aux étoiles les plus faibles visibles à l'œil nu.

Beaucoup plus récemment – en 1856, pour être exact – cette échelle fut rendue quantitative par l'astronome anglais Norman Robert Pogson. Celui-ci montra qu'une étoile de première magnitude était en moyenne cent fois plus brillante qu'une étoile de sixième magnitude. Pogson proposa donc de faire correspondre une différence d'une magnitude à un rapport constant (égal à 2,512) des luminosités. Ainsi, une étoile de quatrième magnitude est 2,512 fois plus brillante qu'une étoile de cinquième magnitude, et 2,512 × 2,512 (soit environ 6,3) fois plus brillante qu'une étoile de sixième magnitude.

Dans cette échelle, 61 Cygni est une étoile faible, de magnitude 5,0 (les méthodes astronomiques modernes permettent d'apprécier le dixième et parfois le centième de magnitude). Capella est au contraire une des plus brillantes, avec une magnitude de 0,9. Alpha Centauri est encore plus brillante : sa magnitude vaut 0,1. Pour les étoiles les plus brillantes, l'échelle continue à travers la magnitude zéro jusqu'aux magnitudes négatives. Sirius, la plus brillante des étoiles de notre ciel, a une magnitude de – 1,42. La planète Vénus atteint – 4,2, la pleine lune – 12,7, le Soleil – 26,9.

Ce sont là des *magnitudes apparentes*, qui caractérisent les étoiles telles que nous les voyons, et qui dépendent à la fois de leur éclat absolu et de leur distance. Les astronomes fondent l'échelle des *magnitudes absolues* sur l'éclat qu'aurait la même étoile ramenée à une distance de 10 parsecs, ou 32,6 années-lumière (le *parsec* est la distance à laquelle une étoile aurait une parallaxe d'une seconde d'arc, soit une distance de 3,26 années-lumière).

Bien que Capella semble plus faible que Sirius et Alpha Centauri, c'est en réalité une source de lumière beaucoup plus puissante ; seulement, elle est beaucoup plus loin de nous. Si les trois étoiles étaient ramenées à la distance standard, Capella serait de beaucoup la plus brillante : sa magnitude absolue vaut – 0,1, celle de Sirius 1,3 et celle d'Alpha Centauri 4,8. C'est aussi à peu près la magnitude absolue de notre Soleil (4,86), qui est donc une étoile moyenne.

Revenons maintenant aux céphéides. En 1912, Henrietta Leavitt, astronome de l'observatoire de Harvard, étudia le plus petit des deux Nuages de Magellan – deux énormes amas d'étoiles de l'hémisphère sud, observés pour la première fois lors du tour du monde de Ferdinand Magellan. Parmi les étoiles du Petit Nuage de Magellan, Mlle Leavitt détecta vingt-cinq céphéides. Elle enregistra la période de variation de chacune d'elles, et constata avec surprise que cette période était d'autant plus longue que l'étoile était plus brillante.

Comme ce n'est pas le cas pour les céphéides de notre voisinage, pourquoi serait-ce le cas pour celles du Petit Nuage de Magellan ? Dans notre voisinage, nous ne connaissons que les magnitudes apparentes des céphéides. Ne connaissant ni leur distance, ni leur luminosité absolue, nous ne pouvons pas établir de relation entre celle-ci et la période. Mais dans le Petit Nuage de Magellan, toutes les étoiles sont pratiquement à la même distance de nous, car nous sommes extrêmement loin du Nuage lui-même. C'est comme si un New-Yorkais tentait d'évaluer les distances qui le séparent de divers habitants de Chicago : il trouverait qu'ils sont tous à la même distance de lui – car par rapport à une distance moyenne de l'ordre de mille kilomètres, quelques kilomètres de plus ou de moins ne jouent pas un grand rôle. De la même façon, une étoile située à un bout du Nuage est à peu près à la même distance de nous qu'une étoile située à l'autre bout.

Toutes les étoiles du Petit Nuage de Magellan étant à peu près à la même distance de nous, leur magnitude apparente reflète en valeur relative leur magnitude absolue. Aussi Henrietta Leavitt considéra-t-elle comme vraie en réalité la relation apparente qu'elle y décela : la période des céphéides augmente de façon continue avec leur magnitude absolue. Elle put donc tracer une *courbe période-luminosité* – une courbe montrant quelle période doit avoir une céphéide de magnitude absolue donnée, et réciproquement quelle magnitude absolue doit avoir une céphéide de période donnée.

Si toutes les céphéides de l'Univers se comportaient comme celles du Petit Nuage de Magellan (hypothèse qui semble raisonnable), cela donnerait aux astronomes un moyen d'évaluer les distances relatives aussi loin que l'on pouvait détecter des céphéides dans les meilleurs télescopes. Si on trouvait deux céphéides de même période, on pouvait admettre qu'elles avaient la même magnitude absolue. Dès lors, leurs magnitudes apparentes permettaient de calculer le rapport de leurs distances : si l'une nous paraissait quatre fois plus brillante que l'autre, cela voulait dire qu'elle était deux fois plus proche. On put donc ainsi représenter à l'échelle l'ensemble des céphéides, et pour connaître de façon absolue leurs distances, il suffisait de pouvoir mesurer l'une d'entre elles.

Malheureusement, même la céphéide la plus proche de nous, l'Étoile polaire, est à des centaines d'années-lumière, beaucoup trop loin pour que l'on puisse mesurer sa distance par la méthode des parallaxes. Aussi, les astronomes durent-ils utiliser des méthodes indirectes, comme par exemple le mouvement propre. (Souvenons-nous que c'est le grand mouvement propre de 61 Cygni qui avait conduit Bessel à voir là une étoile relativement proche.) Il fallut divers dispositifs de mesure de ce mouvement propre et des méthodes statistiques d'exploitation des mesures. La procédure était assez compliquée, mais permit d'obtenir les distances approchées d'un certain nombre de groupes d'étoiles contenant des céphéides. A partir de ces distances et des magnitudes apparentes de ces céphéides, on put déterminer leurs magnitudes absolues et – connaissant leurs périodes – porter enfin des unités sur les axes de la courbe d'Henrietta Leavitt.

En 1913, l'astronome danois Ejnar Hertzsprung détermina qu'une céphéide de magnitude absolue – 2,3 avait une période de 6,6 jours. A partir de là, la courbe d'Henrietta Leavitt permit de déterminer la magnitude absolue – et donc la distance – de n'importe quelle céphéide. (Par ailleurs, on s'aperçut que ces magnitudes absolues étaient celles de grosses étoiles brillantes, beaucoup plus lumineuses que notre Soleil. Les variations de leur éclat sont probablement le résultat d'une sorte de pulsation, l'étoile semblant alternativement se gonfler et se rétrécir, dans une respiration gigantesque...)

Quelques années plus tard, l'astronome américain Harlow Shapley recommença le travail et trouva pour une céphéide de magnitude absolue – 2,3 une période de 5,96 jours. L'accord entre ces deux résultats était suffisamment bon pour permettre aux astronomes d'aller de l'avant : ils disposaient enfin du « mètre » qui allait leur permettre de mesurer les distances entre étoiles.

LA TAILLE DE LA GALAXIE

En 1918, Shapley commença à évaluer les dimensions de la Galaxie à l'aide des céphéides qui en font partie. Il s'intéressa surtout aux céphéides

qui font partie des groupes d'étoiles connus sous le nom d'*amas globulaires* – des groupes denses, de forme sphérique, regroupant entre dix mille étoiles et dix millions, pour des diamètres de l'ordre de la centaine d'années-lumière.

Ces amas (dont la nature a été déterminée vers 1800 par Herschel) constituent un environnement tout à fait différent de celui qui caractérise notre région de l'espace. Au centre des amas les plus importants, les étoiles sont « entassées » à 50 par parsec cube, au lieu d'une étoile pour dix parsecs cubes comme dans notre voisinage. Dans ces conditions, la lumière du ciel étoilé est beaucoup plus intense que celle de notre pleine lune, et une planète hypothétique située vers le centre de l'un de ces amas ne connaîtrait pas de nuit véritable.

On connaît environ cent amas globulaires dans notre galaxie, et il y en a probablement autant qui n'ont pas encore été détectés. Shapley constata que leurs distances, par rapport à nous, allaient de 20 000 à 200 000 années-lumière. (L'amas le plus proche, dans la direction de la constellation du Centaure, est visible à l'œil nu, comme une « étoile », qui a reçu le nom d'Omega Centauri. Le plus lointain, NGC 2419, est si loin qu'il fait à peine partie de la Galaxie.)

Shapley observa que ces amas globulaires formaient une immense sphère, que le plan de la Galaxie coupait en deux moitiés, et qu'ils entouraient la région centrale de la Galaxie comme un halo. Il fit l'hypothèse, assez naturelle, que cette sphère avait le même centre que la Galaxie. Il calcula ainsi la position du centre de la sphère : dans la direction de la constellation du Sagittaire (qui est bien dans la Voie lactée) et à environ 50 000 années-lumière de nous. Ainsi, notre Soleil, loin d'être proche du centre de la Galaxie (comme l'avaient supposé Herschel et Kapteyn) occupe une position tout à fait excentrée, près du bord.

Dans le modèle de Shapley, la Galaxie était une gigantesque lentille dont le diamètre valait à peu près 300 000 années-lumière. Cette fois, c'était une estimation trop forte, comme une autre méthode de mesure n'allait pas tarder à le montrer.

Puisque la Galaxie avait la forme d'un disque, les astronomes imaginaient volontiers, depuis William Herschel, qu'elle tournait dans l'espace. En 1926, l'astronome hollandais Jan Oort se proposa de mesurer cette rotation. La Galaxie n'étant pas un corps solide, mais un ensemble de nombreuses étoiles isolées, on ne doit pas s'attendre à ce qu'elle tourne d'un seul bloc, comme une roue. Au contraire, les étoiles proches du centre de masse du disque doivent tourner autour de celui-ci plus vite que celles qui en sont plus éloignées (de la même façon que les planètes les plus proches du Soleil mettent moins de temps à parcourir leur orbite). Par conséquent, les étoiles situées vers le centre de la Galaxie (c'est-à-dire dans la direction du Sagittaire) devraient tourner plus vite que notre Soleil, tandis que celles qui sont plus éloignées de ce centre (et qui se trouvent dans la direction de la constellation des Gémeaux) devraient prendre peu à peu du retard par rapport à nous dans leur révolution. Et plus une étoile est lointaine, plus cette différence devrait être importante.

A partir de ces hypothèses, on pouvait se servir des déplacements relatifs des étoiles pour calculer la vitesse de rotation autour du centre de la Galaxie. On trouva ainsi que le Soleil et ses voisines se déplaçaient à une vitesse de l'ordre de 240 km/s par rapport au centre de la Galaxie, et mettaient à peu près 200 millions d'années à en faire le tour. (L'orbite

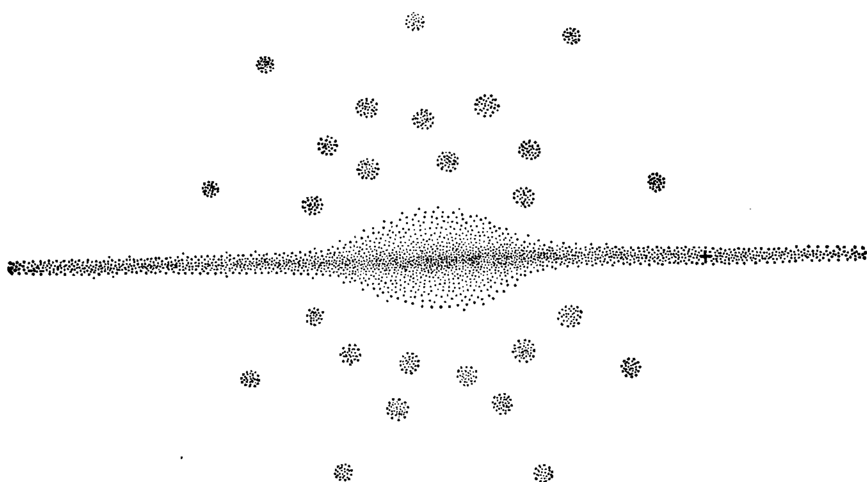
du Soleil est presque circulaire, mais celle de certaines étoiles, Arcturus par exemple, est nettement elliptique. C'est parce que les étoiles n'ont pas des orbites parfaitement circulaires que le Soleil semble se déplacer vers la constellation de la Lyre.)

Ayant estimé la vitesse de rotation, les astronomes purent en déduire l'intensité du champ gravitationnel du centre de la Galaxie, et par conséquent la masse de celle-ci. La région centrale de la Galaxie (qui contient la plus grande partie de sa masse) se révèle ainsi plus de cent milliards de fois plus massive que notre Soleil. Celui-ci étant une étoile de taille supérieure à la moyenne, on peut estimer que notre galaxie contient deux ou trois cents milliards d'étoiles – deux ou trois mille fois plus que dans l'estimation d'Herschel.

A partir des orbites décrites par les étoiles, on peut aussi localiser le centre autour duquel elles tournent. Ce centre est bien dans la direction du Sagittaire, comme l'avait estimé Shapley, mais il n'est qu'à 27 000 années-lumière de nous, ce qui donne, pour le diamètre de la Galaxie, une valeur de 100 000 années-lumière, au lieu de 300 000. Suivant ces nouvelles estimations – qui sont encore considérées comme correctes – l'épaisseur du disque est de l'ordre de 20 000 années-lumière vers le centre, et diminue vers le bord : au niveau de notre Soleil, situé aux deux tiers de la distance entre le centre et le bord de la Galaxie, cette épaisseur vaut environ 3 000 années-lumière (figure 2.3). Mais ce ne sont là que des estimations grossières : la Galaxie n'a pas de limites bien définies.

Puisque le Soleil est si proche du bord de la Galaxie, pourquoi la Voie lactée ne paraît-elle pas beaucoup plus brillante dans la direction du centre que dans la direction opposée, celle du bord ? Quand nous regardons dans la direction du Sagittaire, nous sommes tournés vers le gros des effectifs

Figure 2.3. Notre galaxie, représentée schématiquement de profil. Les amas globulaires sont disposés autour de la région centrale de la Galaxie. La position de notre Soleil est indiquée par une croix.



de la Galaxie, quelque deux cents milliards d'étoiles, tandis que dans la direction opposée (vers le bord de la Galaxie), il n'y en a que quelques millions. Pourtant, la bande de la Voie lactée paraît aussi lumineuse dans toutes les directions. C'est sans doute que d'énormes nuages de poussière s'interposent entre nous et le centre de la Galaxie. Il est possible que la moitié de la masse des régions extérieures de la Galaxie soit formée de tels nuages de poussières et de gaz. Selon toute vraisemblance, nous ne percevons que 1/10 000 au plus de la lumière émise par les régions centrales de la Galaxie.

Voilà pourquoi Herschel avait supposé – de même que les premiers astronomes à avoir étudié la question – que le système solaire était au centre de la Galaxie. C'est aussi, semble-t-il, ce qui conduisit Shapley à surestimer la taille de la Galaxie. Certains des amas étudiés étaient rendus moins brillants par les nuages de poussières et de gaz, et les céphéides qu'ils contiennent étaient moins brillantes (et donc plus lointaines...) qu'elles ne l'auraient été en l'absence de nuages.

AGRANDIR L'UNIVERS

Avant même que l'on détermine la taille et la masse de la Galaxie, les céphéides des Nuages de Magellan (parmi lesquelles Leawitt découvrit la relation fondamentale entre la période et la luminosité) permettaient d'estimer la distance qui nous sépare des Nuages : plus de 100 000 années-lumière. Les meilleures mesures actuelles mettent le Grand Nuage de Magellan à 150 000 années-lumière de nous, et le Petit à 170 000 années-lumière. Le Grand Nuage ne dépasse pas la moitié de notre galaxie (en diamètre) et le Petit le cinquième. De plus, les étoiles y semblent plus dispersées que dans la Galaxie. Le Grand Nuage de Magellan contient cinq milliards d'étoiles (au plus un vingtième de l'effectif de notre galaxie), et le Petit seulement un milliard et demi.

Voilà donc la situation au début des années vingt : l'Univers connu mesurait 200 000 années-lumière de diamètre, et se composait de notre galaxie et de ses deux voisins. Existait-il quelque chose en dehors de cela ?

Le doute planait en ce qui concerne certaines taches lumineuses floues, appelées *nébuleuses* (du mot grec signifiant « nuage »), et que les astronomes avaient remarquées depuis longtemps. L'astronome français Charles Messier en catalogua 103 en 1781 (dont beaucoup sont encore désignées par le nombre qu'il leur a assigné, précédé de la lettre *M* pour Messier).

Ces nébulosités étaient-elles simplement les nuages qu'elles semblaient être ? Certaines d'entre elles, comme la Nébuleuse d'Orion (découverte en 1656 par l'astronome hollandais Christiaan Huygens) étaient effectivement des nuages de gaz et de poussières (dont la masse, dans le cas de la Nébuleuse d'Orion, vaut 500 fois celle du Soleil), illuminés par des étoiles chaudes situées à l'intérieur. D'autres nébulosités se révélèrent être des « amas globulaires », groupes d'étoiles très nombreuses.

Mais il restait des taches lumineuses floues qui ne semblaient contenir aucune étoile. Pourquoi donc étaient-elles lumineuses ? En 1845, l'astronome britannique William Parsons (troisième comte de Rosse), muni d'un télescope de 1,80 mètre de diamètre qu'il passa sa vie à construire,

remarqua que certaines de ces taches avaient une structure en spirale, ce qui permettait de les baptiser « nébuleuses spirales », mais n'apportait pas grand-chose à l'explication de leur luminosité.

La plus spectaculaire de ces taches, M 31 (la Nébuleuse d'Andromède, ainsi appelée parce qu'elle se trouve dans la constellation portant ce nom), fut étudiée en 1611 par l'astronome Simon Marius. C'est un ovale allongé, faiblement lumineux, dont la taille apparente est à peu près la moitié de celle de la Lune. Pouvait-elle être formée d'étoiles, si lointaines que même les plus gros télescopes ne pouvaient pas les « séparer » (les distinguer les unes des autres) ? Si c'était le cas, la Nébuleuse d'Andromède devait être prodigieusement loin – et prodigieusement grande pour qu'on la voie à cette distance. (Dès 1755, le philosophe allemand Emmanuel Kant imagina l'existence de telles agglomérations d'étoiles très lointaines, qu'il appela des *univers-îles*.)

En 1910, une discussion animée s'engagea sur la question. L'astronome américano-hollandais Adriaan Van Maanen affirmait que la Nébuleuse d'Andromède tournait sur elle-même à une vitesse mesurable : elle devait donc être près de nous ; en effet, si elle était située en dehors de la Galaxie, elle serait trop éloignée pour manifester une rotation perceptible. Shapley, un ami de Van Maanen, s'appuya sur cet argument pour soutenir que la Nébuleuse d'Andromède faisait partie de notre galaxie.

Ce n'était pas du tout l'avis de l'astronome américain Heber Doust Curtis. Bien que normalement aucune étoile ne soit visible dans la Nébuleuse d'Andromède, de temps en temps on y voit apparaître une étoile extrêmement faible. Pour Curtis, c'étaient des *novae,* des étoiles qui deviennent brusquement mille fois plus brillantes. Dans notre galaxie, ces étoiles brillent d'un vif éclat avant de s'éteindre, mais dans la Nébuleuse d'Andromède on les distingue à peine, même à leur luminosité maximale. Curtis en déduisit que cette faiblesse apparente des novae était due à l'éloignement considérable de la Nébuleuse d'Andromède. Les étoiles ordinaires y étaient trop faibles pour que l'on pût les distinguer et se fondaient en une sorte de brouillard de faible luminosité.

Le 26 avril 1920, un débat retentissant opposa sur la question Curtis et Shapley. Dans l'ensemble, il s'est agi d'un match nul, bien que Curtis se soit révélé un orateur étonnant et qu'il ait présenté un plaidoyer impressionnant en faveur de ses idées.

Quelques années plus tard, toutefois, il était devenu clair que Curtis avait raison. D'abord, les chiffres cités par Van Maanen se révélaient faux. On ne sait pas bien pourquoi, mais même les personnes les plus compétentes peuvent faire des erreurs, et c'est apparemment ce qui était arrivé à Van Maanen.

Ensuite, en 1924, l'astronome américain Edwin Powell Hubble pointa vers la Nébuleuse d'Andromède le nouveau télescope de 2,5 mètres du mont Wilson, en Californie. (Ce télescope avait reçu le nom de *télescope de Hooker,* en l'honneur de John B. Hooker, qui avait donné l'argent nécessaire à sa construction.) Ce puissant instrument résolut en étoiles distinctes des portions du bord extérieur de la Nébuleuse, montrant d'emblée que celle-ci, au moins dans certaines de ses parties, ressemblait à la Voie lactée, et qu'il y avait peut-être du vrai dans cette idée d'univers-îles.

Parmi les étoiles du bord de la Nébuleuse d'Andromède, il y a des variables céphéides. A l'aide de ces instruments de mesure, Hubble

détermina que la Nébuleuse se trouvait environ à un million d'années-lumière ! Ainsi, la Nébuleuse d'Andromède se trouvait très loin à l'extérieur de notre galaxie. Compte tenu de sa distance, sa taille apparente en faisait un gigantesque amas d'étoiles, presque aussi grand que notre galaxie.

D'autres nébuleuses se sont révélées elles aussi être des amas d'étoiles, encore plus lointains que la Nébuleuse d'Andromède. On a dû reconnaître dans ces *nébuleuses extragalactiques* des galaxies – de nouveaux univers qui ramenaient le nôtre à n'être plus qu'un exemple parmi bien d'autres dans l'espace. Une fois de plus, l'Univers s'était agrandi. Plus grand que jamais, il mesurait non plus des centaines de milliers, mais des millions d'années-lumière.

LES GALAXIES SPIRALES

Pendant les années trente, les astronomes se sont débattus avec plusieurs énigmes lancinantes à propos de ces galaxies. Tout d'abord, sur la base de leurs distances estimées, elles étaient toutes en apparence plus petites que la nôtre. Cela semblait une coïncidence curieuse que nous habitions précisément la plus vaste des galaxies existantes. Ensuite, les amas globulaires entourant la Nébuleuse d'Andromède semblaient moitié moins brillants que les nôtres. (Andromède possède à peu près autant d'amas globulaires que notre galaxie, et ils sont disposés d'une manière sphérique autour de son centre. Ce résultat semble montrer que Shapley avait raison d'attribuer une disposition analogue à nos propres amas globulaires. Quelques galaxies sont particulièrement riches en amas globulaires. La galaxie M 87, dans la Vierge, en possède au moins mille.)

Le problème le plus sérieux était que les distances des galaxies semblaient attribuer à l'Univers un âge de deux milliards d'années seulement (pour des raisons que nous verrons plus loin dans ce chapitre). Ce résultat était préoccupant, parce que les géologues attribuaient à la Terre elle-même un âge plus avancé, sur la base de preuves apparemment incontestables.

Le commencement d'une réponse est venu pendant la Deuxième Guerre mondiale, quand l'astronome américain d'origine allemande Walter Baade découvrit que le « mètre » dont on s'était servi pour mesurer les distances des galaxies était faux.

En 1942, profitant du black-out imposé par la guerre à la ville de Los Angeles, black-out qui améliorait les conditions d'observation à l'observatoire du mont Wilson, Baade entreprit une étude détaillée de la Nébuleuse d'Andromède au télescope de 2,50 mètres. Grâce à l'amélioration de la visibilité, il a pu distinguer quelques étoiles des régions intérieures de la galaxie. Il remarqua ainsi immédiatement des différences frappantes entre ces étoiles et celles des régions extérieures de la galaxie. Les étoiles les plus brillantes de l'intérieur étaient rougeâtres, celles de l'extérieur bleuâtres. De plus, les géantes rouges de l'intérieur n'étaient pas, et de loin, aussi brillantes que les géantes bleues des régions frontières : les dernières atteignaient 100 000 fois la luminosité de notre Soleil, tandis que les géantes rouges de l'intérieur étaient cent fois moins brillantes. Enfin, les régions extérieures de la Nébuleuse d'Andromède, celles où brillaient les géantes bleues, étaient riches en poussières, tandis que l'intérieur, avec ses étoiles rouges relativement moins brillantes, en était dépourvu.

Baade en déduisit qu'il y avait deux ensembles d'étoiles, de structure et d'histoire distinctes. Il appela *étoiles de population I* les étoiles bleuâtres périphériques, et *étoiles de population II* les étoiles rougeâtres de l'intérieur. Les étoiles de population I se sont révélées relativement jeunes, et riches en métaux, avec des orbites quasi circulaires autour du centre de la galaxie, orbites situées dans le plan de symétrie de la galaxie. Les étoiles de population II sont relativement vieilles, pauvres en métaux, et leurs orbites sont nettement elliptiques et très inclinées sur le plan de la galaxie. Depuis la découverte de Baade, ces deux groupes d'étoiles ont été subdivisés en plusieurs sous-groupes.

Après la guerre, on installa sur le mont Palomar un nouveau télescope, de cinq mètres de diamètre, appelé télescope Hale, en l'honneur de l'astronome américain George Ellery Hale qui supervisa sa construction. Baade l'utilisa pour poursuivre ses recherches. Il constata certaines irrégularités dans la répartition des deux populations, suivant le type de galaxie observé. Les galaxies dites « elliptiques » (de forme elliptique et de structure interne assez uniforme) semblaient constituées surtout d'étoiles de population II, de même d'ailleurs que les amas globulaires de toutes les galaxies. Au contraire, dans les *galaxies spirales* (munies de « bras » qui leur donnent l'aspect de tourbillons), les bras étaient formés d'étoiles de population I, se détachant sur un fond d'étoiles de population II.

On estime qu'environ 2 % seulement des étoiles de l'Univers sont de population I. Mais notre Soleil et les étoiles familières du voisinage se rangent dans cette catégorie. De ce seul fait, nous pouvons déduire que notre galaxie est une galaxie spirale, et que nous nous trouvons dans un de ses bras (de là viennent les nombreux nuages de poussières, sombres ou lumineux, de notre voisinage : les bras d'une galaxie spirale sont pleins de poussières). Les photographies de la Nébuleuse d'Andromède révèlent qu'elle aussi est du type « galaxie spirale ».

Revenons maintenant au problème des distances. Baade a commencé par comparer les céphéides présentes dans les amas globulaires (population II) et celles qui se trouvent dans notre bras de la Galaxie (population I). Ces deux types de céphéides se sont révélés très différents en ce qui concerne la relation entre la période et la luminosité. Les céphéides de population II suivent la courbe période-luminosité établie par Leavitt et Shapley. C'est à partir de là que Shapley avait mesuré, avec une précision raisonnable, les distances des amas globulaires et les dimensions de notre galaxie. Mais on aperçut alors que des calculs fondés sur les céphéides de population I donneraient des résultats tout différents ! Une céphéide de population I étant quatre ou cinq fois plus lumineuse qu'une céphéide de même période de population II, l'échelle de Leavitt donnerait une estimation fausse de la magnitude absolue. Et si la magnitude absolue est fausse, la distance qu'elle permet de calculer est fausse aussi : l'étoile est beaucoup plus éloignée que ne l'estime un tel calcul.

Hubble avait estimé la distance de la Nébuleuse d'Andromède à partir des céphéides (de population I) de ses bords – les seules étoiles résolues à son époque. Mais, compte tenu des nouveaux résultats, la Nébuleuse est à 2,5 millions d'années-lumière, au lieu d'un peu moins d'un million comme il l'avait pensé. Les autres galaxies ont donc dû être « reculées » dans les mêmes proportions. (Cependant, la Nébuleuse d'Andromède reste quand même une proche voisine : on estime à une vingtaine de millions d'années-lumière la distance moyenne entre galaxies.)

D'un seul coup, la taille de l'Univers connu a plus que doublé, et le problème qui avait obsédé les années trente s'est trouvé résolu. Notre galaxie n'était plus d'une taille supérieure à celles des autres. La galaxie d'Andromède, par exemple, était à coup sûr plus massive que la nôtre. De plus, les amas globulaires de la galaxie d'Andromède paraissaient désormais aussi lumineux que les nôtres : c'est seulement à cause d'une estimation fausse de leur distance qu'ils avaient paru moins lumineux. Enfin, pour des raisons sur lesquelles nous reviendrons bientôt, cette nouvelle estimation des distances donnait à l'Univers un âge nettement plus avancé, en accord avec les estimations des géologues au sujet de l'âge de la Terre.

LES AMAS DE GALAXIES

En doublant les distances des galaxies, on n'en finit pas pour autant avec le problème des dimensions de l'Univers. Il faut maintenant considérer la possibilité de systèmes plus vastes, d'amas de galaxies et d'amas d'amas.

De fait, les télescopes modernes ont montré que les amas de galaxies existaient vraiment. Par exemple, dans la constellation de la Chevelure de Bérénice, il y a un vaste amas ellipsoïdal de galaxies, dont le diamètre vaut à peu près huit millions d'années-lumière. L'amas de la Chevelure contient environ 11 000 galaxies, séparées les unes des autres par une distance moyenne de 300 000 années-lumière (au lieu de trois millions comme cela semble être le cas dans notre voisinage).

Quant à notre galaxie, elle semble faire partie d'un *groupe local,* qui comprend les Nuages de Magellan, la Nébuleuse d'Andromède et les trois petites *galaxies satellites* qui l'accompagnent, plus quelques autres galaxies, formant un effectif total de dix-neuf. Deux de ces galaxies n'ont été découvertes qu'en 1971 par un astronome italien, Paolo Maffei, ce qui leur vaut de s'appeler Maffei I et Maffei II. Cette découverte tardive est due aux nuages de poussières qui nous en séparent.

Dans le groupe local, seules la galaxie d'Andromède, la nôtre et les deux Maffei sont géantes, les autres sont naines. L'une de ces naines, IC 1613, ne contient peut-être que soixante millions d'étoiles, et n'est par conséquent guère plus qu'un gros amas globulaire. Parmi les galaxies, comme parmi les étoiles, les naines sont beaucoup plus nombreuses que les géantes.

Si les galaxies forment des amas, et ces amas des amas, cela veut-il dire que l'Univers continue sans limites, et que l'espace est infini ? Ou y a-t-il une limite, à l'Univers et à l'espace ? Les astronomes semblent être capables de distinguer des objets jusqu'à une dizaine de milliards d'années-lumière, mais là, ils semblent atteindre une limite. Pour comprendre pourquoi, il nous faut maintenant changer de thème de discussion. Après avoir parlé de l'espace, nous allons maintenant nous pencher sur la question du temps.

L'origine de l'Univers

Les mythes proposent de nombreuses versions de la création de l'Univers, plus imaginatives les unes que les autres (et habituellement centrées sur la Terre, le reste étant rapidement évoqué sous le nom de

« ciel »). En général, la date attribuée à la Création n'est pas très lointaine (il faut se dire qu'une période de mille ans paraissait beaucoup plus inconcevable à nos ancêtres qu'un million d'années pour nous).

L'histoire de la Création qui nous est la plus familière est, bien entendu, celle qui figure dans les premiers chapitres de la Genèse. C'est une adaptation des mythes babyloniens, dotée d'une poésie supplémentaire et atteignant à la grandeur morale, au moins d'après certains commentateurs.

On a tenté, à diverses reprises, de déterminer la date de la Création à partir des informations données dans la Bible (règnes des différents rois, temps écoulé entre l'Exode et l'inauguration du temple de Salomon, âge des patriarches avant et après le Déluge). Des érudits israélites du Moyen Age placent la création en 3760 av. J.-C., et le calendrier israélite prend encore cette date pour point de départ. En 1658, l'archevêque anglican James Ussher calcula que la Création avait eu lieu en 4004 av. J.-C. (certains de ses partisans précisent que l'événement a eu lieu à huit heures du soir le 22 octobre de cette année). Certains théologiens de l'Église grecque orthodoxe reculent la date de la Création jusqu'en 5508 av. J.-C.

Jusqu'au XVIII[e] siècle, le monde savant a accepté la version de la Bible, et attribué à l'Univers un âge de six ou sept milliers d'années tout au plus. Le premier à s'élever sérieusement contre cette idée fut le naturaliste écossais James Hutton, dans un livre intitulé *Theory of the Earth* (théorie de la Terre), publié en 1785. Hutton est parti de l'idée que les processus naturels à l'œuvre à la surface de la Terre (formation de montagnes et érosion, creusement des lits de rivière, etc.) s'étaient maintenus au même rythme tout au long de l'histoire de la Terre. Ce *principe d'uniformité* impliquait que ces processus avaient dû se poursuivre pendant très longtemps pour aboutir au résultat actuel. Par conséquent, la Terre devait avoir non pas des milliers, mais des millions d'années.

Les idées de Hutton ont tout de suite été tournées en dérision. Mais la graine germait. Au début des années 1830, le géologue anglais Charles Lyell a repris à son compte les idées de Hutton, dans un ouvrage en trois volumes intitulé *Principes de géologie,* avec tant de clarté et de force que le monde scientifique en a été convaincu. On peut dire que cet ouvrage a été le point de départ de la géologie moderne.

L'ÂGE DE LA TERRE

On a tenté de calculer l'âge de la Terre à partir du principe d'uniformité. Par exemple, connaissant la vitesse de dépôt des sédiments (de l'ordre d'un mètre en 2 500 ans, selon les estimations modernes), on peut calculer l'âge d'une couche sédimentaire à partir de son épaisseur. Mais on s'est vite rendu compte que cette méthode ne permettrait pas une évaluation précise de l'âge de la Terre, à cause des perturbations apportées aux roches par l'érosion, les éboulements, les soulèvements, et d'autres forces. Néanmoins, malgré ses imperfections, la méthode permettait d'attribuer à la Terre un âge d'au moins cinq cents millions d'années.

On a aussi évalué l'âge de la Terre en mesurant l'accumulation du sel par les océans, comme l'avait suggéré le premier Edmund Halley en 1715. Les rivières apportant continuellement du sel à la mer, et l'évaporation

ne lui prenant que de l'eau douce, la concentration de l'Océan en sel devait s'élever. En supposant l'Océan initialement formé d'eau douce, le temps nécessaire pour que les rivières l'amènent à sa concentration saline actuelle de 3 % pouvait atteindre un milliard d'années.

Ce grand âge plaisait beaucoup aux biologistes qui, durant la seconde moitié du XIX[e] siècle, ont essayé de suivre la lente évolution des êtres vivants depuis les organismes unicellulaires jusqu'aux animaux supérieurs. Cette évolution demandait beaucoup de temps, et un milliard d'années paraissait une durée raisonnable.

Pourtant, vers le milieu du XIX[e] siècle, de nouvelles complications sont nées brusquement de considérations astronomiques. Par exemple, le principe de *conservation de l'énergie* soulevait un problème intéressant à propos du Soleil. Celui-ci déversait de l'énergie en quantité colossale, et cela depuis le début des temps accessibles à l'histoire. Si la Terre avait existé depuis un temps considérable, d'où était venue toute cette énergie ? Certainement pas des sources connues. Si le Soleil avait été à l'origine un bloc de charbon brûlant dans une atmosphère d'oxygène, il aurait été entièrement converti en gaz carbonique (à la vitesse à laquelle il produisait de l'énergie) en quelque 2 500 ans.

Le physicien allemand Hermann Ludwig Ferdinand von Helmholtz, l'un des premiers à énoncer la loi de conservation de l'énergie, s'intéressait particulièrement au problème du Soleil. En 1854, il fit remarquer que si le Soleil se contractait, sa masse gagnerait de l'énergie en tombant vers son centre de masse, comme une pierre gagne de l'énergie en tombant. Cette énergie pouvait être convertie en rayonnement. D'après les calculs d'Helmholtz, une contraction de 1/10 000 du rayon du Soleil pouvait lui fournir deux mille ans d'énergie.

Le physicien anglais William Thomson (devenu plus tard Lord Kelvin) approfondit la question et détermina que, sur cette base, la Terre ne pouvait pas avoir plus de cinquante millions d'années ; compte tenu du débit d'énergie du Soleil, il devait s'être contracté à partir d'une taille gigantesque, aussi grande au départ que l'orbite de la Terre (ceci supposant, bien sûr, que Vénus et surtout Mercure étaient plus jeunes que la Terre). Puis Lord Kelvin montra que si la Terre elle-même était fondue à sa naissance, il lui avait fallu environ vingt millions d'années pour se refroidir jusqu'à sa température actuelle.

Vers 1890, deux armées apparemment invincibles campaient sur leurs positions : les physiciens semblaient avoir montré de façon satisfaisante que la Terre ne pouvait être solide depuis plus de quelques millions d'années, tandis que les géologues et les biologistes avaient prouvé, de façon tout aussi satisfaisante, qu'elle devait être solide depuis au moins un milliard d'années.

C'est alors que se produisit un événement nouveau, complètement inattendu, et que les arguments des physiciens commencèrent à s'effondrer.

En 1896, la découverte de la radioactivité montra que l'uranium de la Terre, et les autres substances radioactives, libéraient de grandes quantités d'énergie, et cela depuis un temps considérable. Cette découverte rendait vides de sens les calculs de Kelvin, comme le fit remarquer le premier, en 1904, le physicien britannique d'origine néo-zélandaise Ernest Rutherford, au cours d'une conférence à laquelle assistait, sans cacher sa désapprobation, Kelvin lui-même, alors assez âgé.

Il ne sert à rien de chercher à déterminer le temps que mettrait la Terre à se refroidir si on ne tient pas compte de la chaleur constamment produite par les substances radioactives. En en tenant compte, la Terre pourrait bien avoir besoin de milliards, et non pas de millions d'années, pour se refroidir de l'état de masse fondue jusqu'à sa température actuelle. Elle pourrait même s'échauffer au cours du temps !

En fait, la radioactivité a fini par fournir la meilleure estimation de l'âge de la Terre (comme nous le décrirons au chapitre 6), en permettant aux géologues et aux géochimistes de calculer l'âge des roches à partir des quantités de plomb et d'uranium qu'elles contiennent. L'« horloge radioactive » indique maintenant que certaines roches de la Terre ont au moins trois milliards d'années, et il y a toutes les raisons de croire que la Terre est encore plus âgée que cela. On s'accorde maintenant sur un âge de 4,6 milliards d'années pour la Terre sous la forme solide qui est présentement la sienne. Par ailleurs, quelques pierres rapportées de notre voisine la Lune se sont révélées à peu près de cet âge.

LE SOLEIL ET LE SYSTÈME SOLAIRE

Et le Soleil ? La radioactivité, avec les découvertes concernant le noyau atomique, lui fournit une source d'énergie beaucoup plus grande que toutes celles que l'on connaissait auparavant. En 1930, le physicien anglais Sir Arthur Eddington a lancé une nouvelle série de réflexions en émettant l'idée que la température et la pression au centre du Soleil devaient être extrêmement élevées : la température atteindrait quinze millions de degrés. Dans de telles conditions de température et de pression, les noyaux des atomes pourraient réagir les uns sur les autres, chose impossible dans l'environnement beaucoup plus doux que nous connaissons sur Terre. On sait que le Soleil est surtout formé d'hydrogène. Si quatre noyaux d'hydrogène se combinaient (en formant un atome d'hélium), l'énergie libérée serait considérable.

C'est alors qu'en 1938, le physicien américain d'origine allemande Hans Albrecht Bethe trouva deux moyens possibles d'aboutir à cette réaction dans les conditions qui règnent au centre d'étoiles comme le Soleil : l'un de ces moyens mettait en jeu une conversion directe d'hydrogène en hélium, l'autre utilisait comme intermédiaire un atome de carbone. L'un et l'autre de ces processus peuvent se produire dans les étoiles. Dans le cas de notre Soleil, la conversion directe d'hydrogène en hélium semble le plus important des deux mécanismes. Mais dans les deux cas, il y a transformation de masse en énergie. (Einstein, dans la théorie de la relativité restreinte publiée en 1905, avait montré que masse et énergie étaient deux aspects différents de la même chose, et pouvaient se convertir l'une en l'autre. De plus, la conversion d'une masse très petite suffisait à dégager des quantités énormes d'énergie.)

La cadence à laquelle le Soleil émet de l'énergie implique qu'il perde de la masse à raison de 4 200 000 tonnes par seconde. A première vue, cela semble une perte considérable. Mais la masse du Soleil est de 2 200 000 000 000 000 000 000 000 000 tonnes, ce qui fait qu'il ne perd par seconde que 0,000 000 000 000 000 000 02 % de sa masse. Si le Soleil a cinq milliards d'années, comme le pensent les astronomes, et s'il a émis de l'énergie à la cadence actuelle pendant tout ce temps, il n'a perdu que 1/33 000 de sa masse. On voit donc facilement qu'il peut encore continuer longtemps à émettre de l'énergie comme maintenant.

Ainsi, vers 1940, il semblait raisonnable d'attribuer au système solaire dans son ensemble un âge de cinq milliards d'années environ. Tout le problème de l'âge de l'Univers aurait pu être réglé, si les astronomes n'étaient pas venus mettre à nouveau des bâtons dans les roues : l'Univers semblait trop jeune par rapport au système solaire. Les ennuis venaient de l'observation des galaxies lointaines et d'un phénomène découvert en 1842 par un physicien autrichien nommé Christian Johann Doppler.

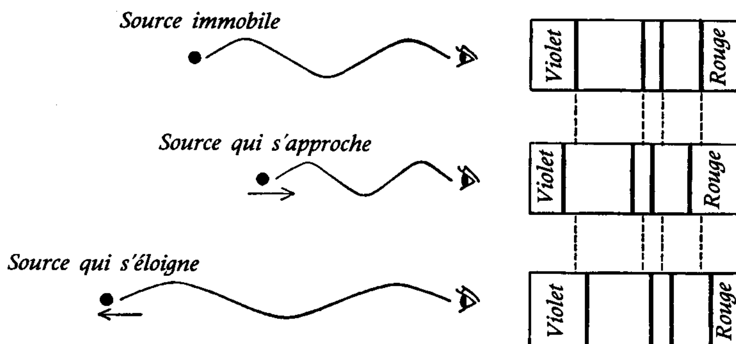
L'*effet Doppler* se manifeste couramment : c'est lui qui fait brusquement baisser la note émise par un sifflet de locomotive, quand celle-ci passe devant nous. Ce changement de fréquence sonore est dû simplement au changement du nombre des ondes sonores qui viennent frapper le tympan en une seconde, changement dû au mouvement de la source.

Comme l'avait suggéré Doppler, cet effet s'applique aussi bien aux ondes lumineuses qu'au son. Quand la lumière émise par une source en mouvement atteint l'œil, il y a un décalage en fréquence – c'est-à-dire en couleur – si la source se déplace assez vite. Par exemple, si la source se rapproche de nous, le nombre des ondes lumineuses reçues par seconde augmente, et la lumière perçue se décale vers les hautes fréquences, vers l'extrémité violette du spectre visible. Au contraire, si la source s'éloigne, il arrive moins d'ondes chaque seconde, et la lumière se décale vers les basses fréquences, vers l'extrémité rouge du spectre.

Les astronomes étudient depuis longtemps le spectre des étoiles, et connaissent bien son aspect normal – un système de raies lumineuses sur fond sombre, ou de raies sombres sur fond lumineux, montrant l'émission ou l'absorption par les atomes de la lumière de certaines longueurs d'onde (de certaines couleurs). Ils ont pu calculer la vitesse d'étoiles qui se déplacent vers nous ou en sens inverse (*vitesse radiale*) en mesurant le décalage vers le violet ou vers le rouge des raies du spectre normal.

C'est le physicien français Armand Louis Hippolyte Fizeau qui a montré, en 1848, que la meilleure manière d'observer l'effet Doppler sur la lumière était de mesurer le décalage des lignes spectrales. Pour cette raison, quand il s'agit de lumière, l'effet Doppler s'appelle *effet Doppler-Fizeau* (figure 2.4).

Figure 2.4. L'effet Doppler-Fizeau. Les raies du spectre se décalent vers l'extrémité violette du spectre (vers la gauche) quand la source lumineuse se rapproche. Quand elle s'éloigne, les raies spectrales se décalent vers l'extrémité rouge (vers la droite).



L'effet Doppler-Fizeau a des applications très variées. Dans le système solaire, il a fourni une nouvelle démonstration de la rotation du Soleil. Les raies spectrales de la lumière provenant du bord du Soleil qui s'approche de nous sont décalées vers le violet, tandis que celles de la lumière provenant de l'autre bord sont décalées vers le rouge, puisque ce bord s'éloigne de nous.

Bien sûr, l'observation des mouvements des taches solaires est un moyen plus évident et plus efficace de mesurer la rotation du Soleil (qui fait environ un tour en vingt-six jours par rapport aux étoiles). Mais l'effet peut aussi servir à mesurer la rotation d'objets sans marque distinctive comme les anneaux de Saturne.

L'effet Doppler-Fizeau peut être utilisé pour des objets situés à n'importe quelle distance, pourvu que ces objets soient assez lumineux pour produire un spectre qui puisse être étudié. Aussi, c'est avec les étoiles que la méthode a connu ses succès les plus spectaculaires.

En 1868, l'astronome britannique Sir William Huggins a mesuré la vitesse radiale de Sirius, et trouvé que cette étoile s'éloignait de nous à une vitesse de 46 km/s (ce qui est très proche des valeurs plus précises calculées par la suite). En 1890, l'astronome américain James Edward Keeler, utilisant des instruments plus précis, obtenait une foule de résultats dignes de confiance ; il a montré, par exemple, qu'Arcturus s'approchait de nous à raison de 6 km/s.

Le même effet permet aussi de déterminer l'existence de systèmes d'étoiles dont le télescope ne peut pas révéler la structure. En 1782, par exemple, un astronome anglais, John Goodricke (sourd-muet mort à vingt-deux ans, un cerveau hors du commun dans un corps tragiquement inadéquat) a étudié l'étoile Algol, dont la luminosité croît et décroît périodiquement. Goodricke a expliqué ce phénomène en imaginant qu'un compagnon sombre tournait autour d'Algol, passant périodiquement devant elle, l'éclipsant et atténuant sa lumière.

Un siècle devait s'écouler avant que cette hypothèse plausible soit confirmée par des résultats expérimentaux. En 1889, l'astronome américain Hermann Carl Vogel montra que les raies du spectre d'Algol se décalaient tantôt vers le rouge, tantôt vers le violet, à la même fréquence que ses variations de luminosité. En effet, l'étoile brillante s'éloigne pendant que le compagnon sombre s'approche, et inversement : Algol apparaissait bien comme une *binaire à éclipses.*

En 1890, Vogel fit une découverte analogue, mais plus générale. Il s'aperçut que certaines étoiles avançaient et reculaient en même temps, c'est-à-dire que leurs raies spectrales se dédoublaient, l'une des composantes étant décalée vers le violet et l'autre vers le rouge. Vogel en conclut qu'il s'agissait de binaires à éclipses, les deux étoiles (brillantes toutes deux) étant si proches l'une de l'autre qu'elles apparaissaient comme une seule dans les meilleurs télescopes. On appelle ces étoiles *binaires spectroscopiques.*

Il n'y avait aucune raison de restreindre aux étoiles de notre galaxie l'utilité de l'effet Doppler-Fizeau. On pouvait aussi s'en servir pour étudier des objets situés hors de la Voie lactée. En 1912, l'astronome américain Vesto Melvin Slipher établit, en mesurant la vitesse radiale de la Nébuleuse d'Andromède, qu'elle s'approchait de nous à 200 km/s. Mais quand il tourna son attention vers d'autres galaxies, il s'aperçut que la plupart s'éloignaient de nous. En 1914, il avait des résultats concernant quinze galaxies, dont treize s'éloignaient de nous à des vitesses atteignant plusieurs centaines de kilomètres par seconde.

A mesure que la recherche progressait dans ce domaine, la situation devenait de plus en plus remarquable. A part quelques-unes des plus proches, toutes les galaxies s'éloignaient de nous. De plus, à mesure que les progrès techniques permettaient d'étudier des galaxies de plus en plus faibles – et donc probablement de plus en plus lointaines –, leur décalage vers le rouge devenait de plus en plus important.

En 1929, Hubble, de l'observatoire du mont Wilson, suggéra que les vitesses auxquelles les galaxies s'éloignaient de nous étaient proportionnelles à leurs éloignements respectifs. Si la galaxie A était deux fois plus éloignée de nous que la galaxie B, alors A s'éloignait deux fois plus vite que B. C'est ce que l'on appelle parfois la *loi de Hubble.*

Les observations ont continué à confirmer cette loi. A partir de 1929, Milton La Salle Humason, au mont Wilson, a utilisé le télescope de 2,50 mètres pour obtenir les spectres de galaxies encore plus faibles. Les plus lointaines qu'il ait réussi à analyser s'éloignaient à une vitesse de 40 000 km/s. Quand le télescope de 5 mètres entra en service, on put étudier des galaxies encore plus lointaines, et vers 1860, on détecta des objets si lointains que leurs vitesses d'éloignement atteignaient 240 000 km/s.

A quoi cela peut-il être dû ? Eh bien, imaginez un ballon avec des points de couleurs peints à sa surface. Quand on gonfle le ballon, les points s'écartent. Un lutin minuscule debout sur l'un des points verrait tous les autres s'éloigner, et d'autant plus vite qu'ils sont plus lointains. Cet effet serait le même quel que soit le point sur lequel le lutin se tiendrait.

Les galaxies se comportent comme si l'Univers grandissait à la manière de la surface tridimensionnelle d'un ballon à quatre dimensions. Les astronomes, de nos jours, ont en général admis cette expansion, et les « équations du champ » d'Einstein, dans sa théorie de la relativité générale, peuvent être rendues compatibles avec un Univers en expansion.

LE BIG BANG

Si l'Univers a constamment grandi, il est logique de supposer qu'il était plus petit dans le passé qu'il ne l'est maintenant ; et qu'à un certain moment, il y a très longtemps, il a débuté sous la forme d'un noyau de matière dense.

Le premier à suggérer cette possibilité, en 1922, a été le mathématicien russe Alexandre Alexandrovitch Friedmann. Les résultats relatifs à l'éloignement des galaxies n'avaient pas encore été présentés par Hubble, et Friedmann s'appuyait seulement sur la théorie, à partir des équations d'Einstein. Mais Friedmann est mort de la typhoïde trois ans plus tard, à l'âge de trente-sept ans, et ses travaux sont restés pratiquement inconnus.

En 1927, l'astronome belge Georges Lemaître, sans connaître apparemment les travaux de Friedmann, est arrivé à une représentation analogue de l'Univers en expansion. Puisqu'il était en expansion, l'Univers avait dû, dans le passé, être très petit, et aussi dense que possible. Lemaître a donné à cet état le nom d'*œuf cosmique.* Suivant les équations d'Einstein, il ne pouvait que grandir ; et compte tenu de sa densité énorme, cette expansion devait avoir lieu avec une violence prodigieuse. Les galaxies actuelles sont des fragments de l'œuf cosmique, et leur mouvement d'éloignement est l'écho de cette très lointaine explosion.

Les travaux de Lemaître, à leur tour, sont passés inaperçus, jusqu'à ce que l'astronome anglais plus connu, Arthur Stanley Eddington, les signale à l'attention du monde scientifique.

Mais c'est un physicien américain d'origine russe, George Gamow, qui a réellement répandu, dans les années trente et quarante, cette idée d'un début explosif de l'Univers. Il a donné à cette explosion initiale le nom de *big bang,* sous lequel elle est maintenant désignée dans le monde entier.

Tout le monde n'était pas d'accord avec l'idée que le big bang constituait le point de départ de l'Univers. En 1948, deux astronomes d'origine autrichienne, Hermann Bondi et Thomas Gold, ont présenté une théorie – qui devait être étendue et popularisée plus tard par l'astronome anglais Fred Hoyle – acceptant l'expansion de l'Univers mais refusant le big bang. A mesure que les galaxies s'écartent les unes des autres, de nouvelles galaxies se forment entre elles, leur matière étant créée à partir de rien, à une cadence trop faible pour être détectable avec les techniques actuelles. En conséquence, l'Univers reste essentiellement le même de toute éternité. Il avait son aspect actuel il y a des milliards et des milliards d'années, et continuera d'avoir le même aspect pendant des milliards d'années, ce qui fait qu'il n'a ni début ni fin. Cette théorie, connue sous le nom de *création continue,* a pour résultat un *Univers stable.*

Pendant plus de dix ans, la controverse a été très animée entre les partisans du big bang et ceux de la création continue, sans qu'aucun résultat expérimental permette de trancher le débat.

En 1949, Gamow avait fait remarquer que, dans l'hypothèse du big bang, le rayonnement émis lors de cet événement avait dû perdre de l'énergie au fur et à mesure de l'expansion de l'Univers, et devait maintenant exister sous forme d'ondes radio venant de toutes les parties du ciel et constituant un fond continu. Ce rayonnement devait être caractéristique d'objets à une température voisine de 5 K (cinq degrés au-dessus du zéro absolu, soit – 268 °C). Cette idée a été reprise et développée par le physicien américain Robert Henry Dicke.

En mai 1964, un physicien américain d'origine allemande, Arno Allan Penzias, et un radioastronome américain, Robert Woodrow Wilson, suivant les indications de Dicke, ont détecté un fond continu d'ondes radio dont les caractéristiques étaient proches de celles qu'avait prédites Gamow. Cette radiation attribuait à l'Univers une température moyenne de 3 K.

Pour beaucoup d'astronomes, cette découverte du fond continu permet de considérer comme valable la théorie du big bang. On admet généralement maintenant qu'il y a bien eu un big bang, et on a abandonné l'idée de création continue.

Quand le big bang s'est-il produit ?

Grâce au décalage vers le rouge, qui est facile à mesurer, nous connaissons avec une bonne certitude la vitesse à laquelle les galaxies s'éloignent. Il nous faudrait connaître aussi la distance à laquelle elles se trouvent. Plus elles sont lointaines, plus il leur a fallu de temps pour atteindre leur position actuelle, à la vitesse à laquelle elles s'écartent. Mais ce n'est pas facile de déterminer ces distances.

On considère généralement comme plausible une estimation à quinze milliards d'années de l'ancienneté du big bang. Si un *éon* représente un milliard d'années, le big bang se serait produit il y a quinze éons, mais c'est aussi possible qu'il se soit produit il y a vingt éons, ou seulement dix.

Et avant ? Que s'est-il passé avant le big bang ? D'où vient l'œuf cosmique ?

Certains astronomes imaginent qu'en réalité l'Univers a commencé sous la forme d'un gaz très ténu, qui se serait lentement condensé, en formant peut-être des étoiles et des galaxies, et aurait continué à se contracter, jusqu'à former un œuf cosmique au cours d'un *big crunch*. Sitôt formé, l'œuf cosmique aurait explosé (big bang) en formant à nouveau des étoiles et des galaxies, mais cette fois en expansion, jusqu'à redevenir, dans un futur lointain, un gaz ténu.

Si nous pouvions voir dans l'avenir, peut-être verrions-nous l'Univers se dilater indéfiniment, devenant de plus en plus ténu, avec une densité moyenne de plus en plus faible, et s'approchant de plus en plus du vide du néant. Et si nous pouvions voir dans le passé, au-delà du big bang, en imaginant que le temps marche à l'envers, peut-être verrions-nous de la même façon l'Univers se dilater à l'infini et tendre vers le vide.

Un modèle à « un aller-retour » de ce genre, l'homme occupant une place assez proche du big bang pour que la vie soit possible (sinon, nous ne serions pas là pour observer l'Univers et tenter d'interpréter nos observations) s'appelle un *univers ouvert.*

Nous n'avons pour l'instant aucun moyen (et nous ne l'aurons peut-être jamais) d'obtenir des informations sur ce qui se passait avant le big bang, et certains astronomes se refusent à toute spéculation à ce sujet. Lors de discussions récentes, on a proposé l'idée que l'œuf cosmique se serait formé à partir de rien, ce qui donnerait un modèle à « un aller-simple », et non plus un aller-retour. Ce serait encore un univers ouvert.

Dans cette hypothèse, il se pourrait que, dans l'infini du néant, un nombre infini de big bangs puisse se produire à des moments divers, et que par conséquent notre Univers ne soit que l'un parmi une infinité d'autres, chacun ayant sa propre masse, son propre degré de développement et, pourquoi pas, son propre système de lois physiques. Peut-être faut-il une combinaison exceptionnelle de lois physiques pour permettre l'apparition d'étoiles, de galaxies et d'êtres vivants, et nous serions alors dans une telle situation exceptionnelle simplement parce que nous ne pourrions être dans aucune autre.

Il va sans dire qu'aucun résultat expérimental ne permet encore de trancher en faveur de l'apparition de l'œuf cosmique à partir de rien, ou en faveur de l'existence d'une infinité d'univers – et il est bien possible qu'il en soit toujours ainsi. Mais dans quel monde vivrions-nous si les scientifiques n'avaient plus le droit, en l'absence de résultats expérimentaux, de laisser vagabonder leur imagination ?

Et d'ailleurs, sommes-nous même sûrs que l'Univers va se dilater indéfiniment ? Il le fait en dépit de son attraction gravitationnelle, et peut-être cette attraction deviendra-t-elle suffisante pour stopper l'expansion, et provoquer une contraction de l'Univers ? Après son expansion, l'Univers pourrait se contracter jusqu'à s'écraser sur lui-même (big crunch) et disparaître – ou au contraire « rebondir » pour se dilater à nouveau avant de se contracter une nouvelle fois, en une série infinie d'oscillations. Dans les deux cas, ce serait un *univers fermé.*

Il n'est peut-être pas impossible de savoir si l'Univers est fermé ou ouvert : nous reviendrons sur ce sujet au chapitre 7.

La mort du Soleil

L'expansion de l'Univers, même si elle se poursuit indéfiniment, n'affecte pas directement chaque galaxie, ou chaque groupe local de galaxies. Même si toutes les galaxies lointaines finissaient par s'éloigner jusqu'au point où elles seraient hors de portée des meilleurs instruments, notre propre galaxie resterait intacte, ses étoiles solidement maintenues dans son champ gravitationnel. De plus, les autres galaxies du groupe local ne nous quitteraient pas non plus. Néanmoins, des changements à l'intérieur de notre galaxie, indépendants de l'expansion de l'Univers, mais peut-être désastreux pour notre planète et la vie qu'elle porte, ne sont aucunement à exclure.

L'idée même de changements dans les corps célestes est une idée moderne. Les philosophes de la Grèce antique – Aristote en particulier – croyaient le ciel parfait et immuable. Le changement, la corruption et la décrépitude étaient propres à la région imparfaite située à l'intérieur de la plus petite des sphères, celle de la Lune. Cette idée semblait raisonnable, puisque d'une génération à l'autre et même d'un siècle à l'autre, on ne décelait aucun changement important dans le ciel. Bien sûr, de mystérieuses comètes apparaissaient soudain, de temps en temps, fantaisistes dans leurs allées et venues, fantomatiques par le voile qu'elles tendaient devant les étoiles, effrayantes avec leur longue « chevelure » en désordre comme celle de prophétesses de malheur. Vingt-cinq par siècle environ sont visibles à l'œil nu. (Nous parlerons plus en détail des comètes dans le chapitre suivant.)

Pour tenter de réconcilier ces apparitions avec la perfection céleste, Aristote y voyait des phénomènes atmosphériques, appartenant au désordre et à la corruption terrestres. Cette idée devait rester en vigueur jusqu'à la fin du XVIe siècle. Mais en 1577 (avant l'invention de la lunette), l'astronome danois Tycho Brahé tenta de mesurer la parallaxe d'une comète brillante, et s'aperçut qu'elle était trop petite pour être mesurée. Puisqu'on pouvait mesurer celle de la Lune, il fallait donc que la comète soit située bien plus loin qu'elle, et par conséquent il y avait dans le ciel des changements et des imperfections (le philosophe romain Sénèque avait soupçonné cette possibilité au Ier siècle de notre ère).

En réalité, on avait remarqué, bien avant cela, des changements dans les étoiles elles-mêmes, mais apparemment sans que cela ait soulevé une grande curiosité. Par exemple, il y a des étoiles variables qui changent d'éclat d'une nuit à l'autre, même à l'œil nu. Aucun astronome grec n'a jamais signalé d'étoiles variables. Peut-être avons-nous perdu les documents relatifs à ce phénomène, mais peut-être aussi les astronomes grecs avaient-ils simplement choisi de ne pas le voir. Un cas intéressant est celui d'Algol, la deuxième étoile, par ordre d'éclat, dans la constellation de Persée : Algol perd les deux tiers de son éclat, puis les retrouve, les reperd, et ainsi de suite, avec une période de soixante-neuf heures. (Nous savons maintenant, grâce à Goodricke et Vogel, qu'Algol a un compagnon sombre, qui l'éclipse et diminue son éclat avec une période de soixante-neuf heures.) Les astronomes grecs n'ont fait aucune mention d'Algol, pas plus que leurs collègues arabes du Moyen Age. Mais les Grecs plaçaient cette étoile dans la tête de Méduse, démon femelle qui changeait les hommes en pierre, et le nom même d'Algol, en arabe, signifie « vampire »... De

toute évidence, cette étoile bizarre donnait aux Anciens un sentiment de malaise.

Dans la constellation de la Baleine, il y a une étoile (Omicron Baleine) qui varie d'une façon irrégulière. Elle est parfois aussi brillante que l'Étoile polaire, et elle disparaît parfois complètement. Ni les Grecs ni les Arabes n'en ont dit un mot, et le premier à la signaler fut l'astronome hollandais David Fabricius, en 1596. On lui a donné plus tard le nom de Mira (« merveille » en latin) : les astronomes avaient surmonté leur peur des changements dans le ciel.

NOVAE ET SUPERNOVAE

Les Grecs ne pouvaient pas ignorer un phénomène plus remarquable encore : l'apparition dans le ciel de *nouvelles étoiles*. C'est la vue d'une nouvelle étoile, apparue dans la constellation du Scorpion en 134 av. J.-C., qui impressionna Hipparque au point qu'il dessina la première carte du ciel, de manière à faciliter le repérage ultérieur d'événements de ce type.

En l'an 1054, dans la constellation du Taureau, une autre étoile nouvelle apparaissait, d'un éclat prodigieux : elle était plus brillante que Vénus, et pendant des semaines elle fut visible en plein jour ! Les astronomes chinois et japonais relevèrent avec soin sa position, et leurs relevés sont parvenus jusqu'à nous. Quant au monde occidental, le niveau de l'astronomie y était si bas à l'époque qu'il n'existe aucun document relatif à cette étoile, sans doute parce que personne n'a pris la peine de rédiger un tel document.

Les choses avaient changé en 1572, quand une nouvelle étoile, aussi brillante que celle de 1054, apparut dans la constellation de Cassiopée. L'astronomie européenne sortait alors de son long sommeil. Le jeune Tycho Brahé observa avec soin la nouvelle étoile, et publia un livre intitulé *De Nova Stella.* C'est de ce titre que l'on a tiré le nom de *nova* donné à toute étoile nouvelle.

En 1604, une autre nova remarquable apparut dans la constellation du Serpent. Elle n'était pas tout à fait aussi brillante que celle de 1572, mais assez tout de même pour dépasser Mars en éclat. Cette fois, c'est Johannes Kepler qui l'observa, et il laissa un livre à son sujet.

Après l'invention de la lunette, les novae devinrent moins mystérieuses. Ce n'étaient pas du tout de nouvelles étoiles, bien entendu, mais des étoiles faibles devenues soudainement suffisamment brillantes pour être visibles à l'œil nu.

Avec le temps, on en a découvert de plus en plus. Après être devenues, parfois en quelques jours, plusieurs milliers de fois plus brillantes, elles revenaient en quelques mois à leur éclat de départ. On en observait une vingtaine par an dans chaque galaxie, y compris la nôtre.

A partir des décalages Doppler observés pendant la formation des novae, et d'autres détails de leur spectre, il devint manifeste que les novae étaient des étoiles qui explosaient. Dans certains cas, la matière expulsée dans l'espace était visible sous la forme d'une couche de gaz en expansion, illuminée par le reste de l'étoile.

Dans l'ensemble, les novae récentes n'ont pas été particulièrement brillantes. La plus brillante, apparue en juin 1918 dans la constellation de l'Aigle, était, à son maximum, à peu près aussi lumineuse que Sirius,

l'étoile la plus brillante du ciel. Mais aucune nova récente n'a approché l'éclat de Vénus ou de Jupiter, comme celles de Tycho et de Kepler.

La nova la plus remarquable découverte depuis l'invention du télescope n'a pas tout de suite été reconnue comme telle. L'astronome allemand Ernst Hartwig l'a remarquée en 1885, mais même à son maximum elle n'a jamais dépassé la septième magnitude, et n'a jamais été visible à l'œil nu.

Elle était située dans ce que l'on appelait alors la Nébuleuse d'Andromède, et à son maximum elle était dix fois moins lumineuse que l'ensemble de la Nébuleuse. A l'époque, personne ne réalisait l'énorme éloignement de la Nébuleuse d'Andromède, ni ne comprenait qu'il s'agissait en fait d'une galaxie formée de centaines de milliards d'étoiles, ce qui fait que l'éclat de la nova ne souleva aucun étonnement.

Après que Curtis et Hubble eurent calculé la distance de la galaxie d'Andromède, comme on se mit alors à la nommer, l'éclat de la nova de 1885 devint pour les astronomes un sujet de stupéfaction. Les douzaines de novae découvertes dans la galaxie d'Andromède par Curtis et Hubble étaient beaucoup plus faibles que celle de 1885, qui se révélait, compte tenu de sa distance, prodigieusement brillante.

En 1934, l'astronome suisse Fritz Zwicky commença une recherche systématique de novae exceptionnellement brillantes dans les galaxies lointaines. N'importe quelle nova du même ordre que celle de 1885 devait être visible, puisque l'éclat de telles novae approche celui des galaxies tout entières : si la galaxie est visible, la nova aussi. En 1938, il n'en avait pas repéré moins de douze. Il donna à ces novae dont l'éclat approche celui d'une galaxie le nom de *supernovae.* En conséquence, la nova de 1885 reçut enfin son nom de baptême : S Andromède (*S* pour « supernova »).

Tandis que les novae ordinaires atteignent une magnitude absolue de – 8 en moyenne (vues d'une distance de 10 parsecs, elles seraient vingt-cinq fois plus brillantes que Vénus), une supernova pourrait atteindre une magnitude absolue de – 17. Elle serait quatre mille fois plus brillante qu'une nova ordinaire, soit un milliard de fois plus brillante que le Soleil, au moins au moment de son maximum.

Rétrospectivement, on réalise que les novae de 1054, 1572 et 1604 étaient des supernovae, et situées dans notre galaxie : c'est le seul moyen de rendre compte de leur éclat extraordinaire.

Un certain nombre des novae enregistrées par les astronomes chinois de l'Antiquité et du Moyen Age étaient sans doute aussi des supernovae. C'est le cas d'une nova signalée en l'an 185. Et en 1006, dans la constellation très méridionale du Loup, il a dû apparaître une supernova plus brillante que toutes celles des temps historiques : à son maximum, elle a pu atteindre deux cents fois l'éclat de Vénus, soit un dixième de celui de la pleine lune !

Les astronomes, à partir de l'étude des débris qu'il en reste, soupçonnent qu'une supernova encore plus brillante (autant que la pleine lune) est apparue dans la constellation de la Voile, près du pôle Sud, il y a 11 000 ans, alors qu'il n'y avait pas encore d'astronomes pour l'observer, ni d'écriture pour noter son apparition (mais il est possible que certains pictogrammes préhistoriques aient été dessinés pour représenter cet événement).

Le comportement physique des supernovae diffère complètement de celui des novae ordinaires, et les astronomes ont hâte d'étudier en détail le spectre d'une supernova. Le problème est leur rareté : une en cinquante ans en moyenne dans une galaxie donnée. Bien que les astronomes en aient déjà repéré plus de cinquante, elles étaient toutes situées dans des

galaxies lointaines, trop lointaines pour que l'on puisse les étudier en détail. La supernova de 1885 dans la galaxie d'Andromède, la plus proche de nous depuis 350 ans, est apparue une vingtaine d'années avant que la photographie astronomique soit assez au point pour permettre d'enregistrer son spectre.

Mais l'apparition des supernovae se fait au hasard : on en a détecté trois en dix-sept ans dans la même galaxie. Tout espoir n'est pas perdu pour les astronomes terrestres d'observer une supernova dans notre galaxie, et même près de nous. Une étoile en particulier, Eta Carène, est manifestement instable, et devient tantôt plus brillante, tantôt plus faible. En 1840, elle a été pendant un moment la seconde du ciel par ordre d'éclat ! C'est donc une candidate sérieuse, capable de donner d'un moment à l'autre une supernova. Malheureusement, en astronomie, « d'un moment à l'autre » peut aussi bien vouloir dire demain que dans dix mille ans... De plus, la constellation de la Carène, où se trouve cette étoile, est si proche du pôle Sud céleste, comme celles de la Voile et du Loup, que cette supernova ne serait visible ni depuis l'Europe ni depuis la plus grande partie des États-Unis.

Mais quelle est la cause de ces augmentations explosives d'éclat, et pourquoi certaines étoiles donnent-elles des novae et d'autres des supernovae ? La réponse à cette question nous oblige d'abord à une digression.

Dès 1834, Bessel (l'astronome qui devait plus tard mesurer la première parallaxe d'étoile) avait remarqué que Sirius et Procyon avaient des changements de position très légers qui ne correspondaient pas au mouvement de la Terre. Au lieu d'aller tout droit, elles suivaient un chemin onduleux, et Bessel en avait conclu que chacune d'elles devait en fait tourner autour de quelque chose.

A partir du mouvement de Sirius et de Procyon sur leurs orbites, on pouvait calculer dans chaque cas que le « quelque chose » devait avoir un champ gravitationnel intense, et une masse de l'ordre de celle d'une étoile. Le compagnon de Sirius, en particulier, devait être aussi massif que notre Soleil pour rendre compte des mouvements de l'étoile brillante. Aussi les compagnons devaient-ils être des étoiles, mais puisqu'elles étaient invisibles dans les télescopes de l'époque, on les appela *compagnons sombres* ; on pensait que c'étaient des étoiles vieillies, affaiblies par le temps.

Puis en 1862, le fabricant américain d'instruments Alvan Clark, essayant un nouveau télescope, aperçut près de Sirius une étoile faible. Des observations ultérieures ne devaient pas laisser de doutes : il s'agissait bien du compagnon de Sirius. Les deux étoiles tournaient autour de leur centre de masse commun en une cinquantaine d'années. Le compagnon de Sirius (on l'appelle maintenant Sirius B, Sirius lui-même étant Sirius A) n'a qu'une magnitude absolue de 11,2, c'est-à-dire un éclat quatre cents fois moindre que celui du Soleil, bien qu'il ait à peu près la même masse.

Sirius B semblait être une étoile moribonde. Mais en 1914, l'astronome américain Walter Sidney Adams, ayant étudié le spectre de Sirius B, conclut que cette étoile devait être aussi chaude que Sirius A, et plus chaude que le Soleil. Les vibrations atomiques correspondant aux lignes d'absorption trouvées dans son spectre ne pouvaient se produire qu'à des températures très élevées. Mais si cette étoile était si chaude, pourquoi était-elle si peu lumineuse ? La seule réponse possible était qu'elle était beaucoup plus petite que le Soleil. Étant plus chaude, elle émettait plus

de lumière par unité de surface, et pour rendre compte de sa faible émission totale, force était d'admettre qu'elle avait une petite surface. En fait, nous savons maintenant que Sirius B a à peine plus de dix mille kilomètres de diamètre (moins que la Terre !) pour une masse égale à celle du Soleil ! Pour que toute cette masse tienne dans un si petit volume, la densité moyenne de l'étoile doit être 130 000 fois celle du platine.

Ce n'était rien moins qu'un nouvel état de la matière. Heureusement, les physiciens de l'époque en savaient assez pour suggérer immédiatement une explication. Ils savaient que la matière ordinaire est formée de particules minuscules, si minuscules que la plus grande partie du volume d'un atome est de l'« espace vide ». Sous l'effet de pressions énormes, les particules subatomiques peuvent être forcées de se rapprocher, jusqu'à donner une masse superdense. Par ailleurs, même dans la matière superdense de Sirius B, les particules subatomiques sont assez séparées les unes des autres pour pouvoir se déplacer librement, de sorte que cette matière considérablement plus dense que le platine se comporte encore comme un gaz. En 1925, le physicien anglais Ralph Howard Fowler suggéra d'appeler ce type de matière *gaz dégénéré,* et le physicien soviétique Lev Davidovitch Landau fit remarquer dans les années trente que même des étoiles ordinaires comme notre Soleil devaient, dans leur partie centrale, être formées de gaz dégénéré.

Le compagnon de Procyon (Procyon B), découvert en 1896 par J.M. Schaberle à l'observatoire Lick en Californie, s'est aussi révélé être une étoile superdense, avec une masse égale aux cinq huitièmes de celle de Sirius B. Au fil des années, on a trouvé d'autres exemples. On a donné à ces étoiles le nom de *naines blanches,* parce qu'elles sont à la fois petites et assez chaudes pour émettre de la lumière blanche. Les naines blanches sont probablement nombreuses, et formeraient 3 % du nombre total des étoiles. Toutefois, à cause de leur petite taille et de leur faible éclat, on ne peut espérer en découvrir, dans un avenir prévisible, que dans notre voisinage. (Il y a aussi des *naines rouges,* beaucoup plus petites que notre Soleil, mais pas autant que les naines blanches. Les naines rouges sont « froides », et ont une densité normale. Ce sont les plus répandues de toutes les étoiles – sur quatre étoiles, trois sont des naines rouges –, mais à cause de leur faible luminosité, elles sont aussi difficiles à détecter que les naines blanches. Une paire de naines rouges, située à six années-lumière à peine de nous, n'a été découverte qu'en 1948. Des trente-six étoiles connues situées à moins de quatorze années-lumière du Soleil, vingt et une sont des naines rouges, et trois des naines blanches. Il n'y a pas de géantes parmi elles, et deux seulement, Sirius et Procyon, sont plus brillantes que notre Soleil.)

Un an après que Sirius B eut révélé ses étonnantes propriétés, Albert Einstein présenta sa théorie générale de la relativité, qui considérait la gravitation sous un angle nouveau. Les idées d'Einstein permettaient de prédire qu'une lumière émise par une source possédant un champ gravitationnel très intense serait décalée vers le rouge. Adams, passionné par les naines blanches qu'il avait découvertes, étudia avec soin le spectre de Sirius B, et y décela effectivement le décalage vers le rouge prédit par Einstein. C'était non seulement une confirmation de la théorie d'Einstein, mais aussi du caractère superdense de Sirius B, car avec une étoile ordinaire comme notre Soleil, le décalage vers le rouge serait trente fois moins fort. Pourtant, au début des années soixante, on a réussi à déceler ce très petit

décalage prédit par Einstein dans le spectre de notre Soleil, apportant ainsi une nouvelle confirmation à la théorie générale de la relativité.

Mais quel est le rapport entre les naines blanches et les supernovae, sujet qui nous a amenés à cette discussion ? Revenons à la supernova de 1054.

En 1844, le comte de Rosse, fouillant la région du Taureau où les astronomes orientaux avaient signalé la supernova de 1054, y découvrit un petit objet nébuleux. A cause de son irrégularité et de ses « pattes », il donna à cet objet le nom de Nébuleuse du Crabe. Des observations poursuivies pendant des dizaines d'années ont montré que ce nuage de gaz s'étendait lentement. On a pu ainsi mesurer la vitesse apparente de cette expansion, et comme l'effet Doppler-Fizeau permettait de calculer sa vitesse réelle, on a pu en déduire la distance de la Nébuleuse du Crabe : 3 500 années-lumière. A partir de la vitesse d'expansion, on a pu calculer également à quelle époque le gaz avait commencé à s'étendre à partir d'une explosion initiale : le gaz que nous voyons s'est étendu pendant environ neuf cents ans, ce qui correspond bien à la date (1054) à laquelle on a aperçu l'explosion depuis la Terre. Il est donc extrêmement probable que la Nébuleuse du Crabe, qui nous apparaît maintenant comme un nuage de cinq années-lumière environ de diamètre, représente les débris de la supernova de 1054.

Aucune région analogue de gaz turbulent n'a été découverte aux endroits où les supernovae de Tycho et de Kepler avaient été signalées, bien que de petites taches nébuleuses aient été observées à proximité de chacun de ces sites. On connaît toutefois cent cinquante « nébuleuses planétaires », larges anneaux de gaz qui sont peut-être le résultat d'explosions stellaires. Un nuage de gaz particulièrement étendu et ténu, la Nébuleuse de la Dentelle dans le Cygne, est peut-être le reste d'une supernova qui aurait pu être observée sur Terre il y a trente mille ans (plus proche et plus brillante que celle de 1054) s'il y avait eu à cette époque des hommes civilisés.

On a même avancé l'idée que la nébulosité très faible qui enveloppe la constellation d'Orion pourrait résulter d'une supernova encore plus ancienne.

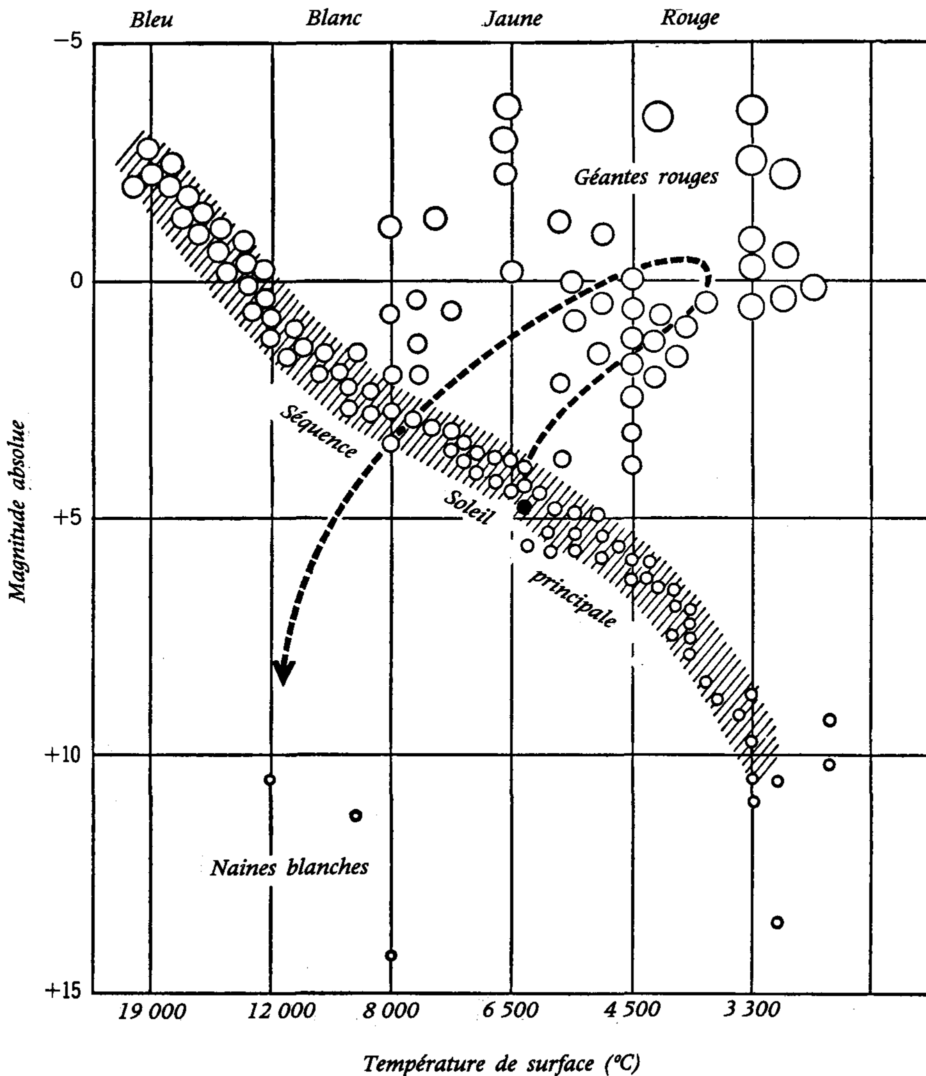
Mais dans tous ces cas, que sont devenues les étoiles qui ont explosé ? Sont-elles simplement disparues dans une énorme bouffée de gaz ? La Nébuleuse du Crabe, par exemple, est-elle *tout* ce qui reste de la supernova de 1054, et va-t-elle s'étendre et se diluer de plus en plus, jusqu'à ce que toute trace de l'étoile disparaisse ? Ou reste-t-il quelque chose qui soit encore une étoile, mais trop petite et trop faible pour être décelable ? Autrement dit, reste-t-il à cet endroit une naine blanche (ou quelque chose d'encore plus bizarre), et les naines blanches seraient-elles les « cadavres », pour ainsi dire, d'étoiles qui ont autrefois été comme notre Soleil ? Ces questions nous amènent au problème de l'évolution des étoiles.

L'ÉVOLUTION DES ÉTOILES

Parmi les étoiles qui sont proches de nous, les brillantes semblent chaudes et les faibles moins chaudes, suivant une courbe assez régulière. Si on représente la température de surface des étoiles en fonction de leur magnitude absolue, la plupart des étoiles familières se regroupent dans

une bande étroite, allant régulièrement du « froid-faible » au « chaud-brillant ». Cette bande s'appelle la *séquence principale.* Elle a été tracée en 1913 par l'astronome américain Henry Norris Russell, reprenant les travaux analogues de Hertzsprung (l'astronome qui avait déterminé le premier les magnitudes absolues des céphéides). Le diagramme montrant la séquence principale s'appelle donc *diagramme de Hertzsprung-Russell,* ou *diagramme H-R* (figure 2.5).

Figure 2.5. Le diagramme de Hertzsprung-Russell. La ligne en pointillés indique l'évolution d'une étoile. Les tailles relatives des étoiles sont schématiques, et ne sont pas à l'échelle.



La séquence principale ne regroupe pas toutes les étoiles. Il y a des étoiles rouges qui, en dépit de leur température de surface plutôt faible, ont de grandes magnitudes absolues, parce que leur matière assez ténue occupe un volume considérable, et que leur émission – assez faible – par unité de surface est multipliée par une surface énorme, donnant une magnitude absolue importante. Parmi ces *géantes rouges,* les plus connues sont Bételgeuse et Antarès. Elles sont si « froides », a-t-on découvert en 1964, que beaucoup d'entre elles ont des atmosphères riches en vapeur d'eau qui, à la température de notre Soleil, serait décomposée en hydrogène et oxygène. Les naines blanches, très chaudes, sont également en dehors de la séquence principale.

En 1924, Eddington a fait remarquer que l'intérieur de chaque étoile devait être très chaud. A cause de la grande masse de l'étoile, sa force gravitationnelle est immense. Si l'étoile ne s'effondre pas sur elle-même, c'est que ces forces gravitationnelles sont compensées par une pression interne très élevée – correspondant à des températures prodigieusement hautes. Plus l'étoile est massive, plus la température centrale doit être élevée pour équilibrer la force gravitationnelle. Pour maintenir cette température élevée et cette pression de radiation, les étoiles plus massives doivent brûler de l'énergie plus vite, et être plus brillantes que celles qui sont moins massives : c'est la *relation masse-luminosité.* Cette relation est très importante, car la luminosité varie comme la sixième ou septième puissance de la masse. Si celle-ci est multipliée par 3, la luminosité s'accroît d'un facteur égal à la sixième ou septième puissance de trois – c'est-à-dire environ 750 fois.

Par conséquent, les étoiles massives dépensent généreusement leur hydrogène, et vivent moins longtemps. Notre Soleil possède suffisamment d'hydrogène pour des milliards d'années d'émission à la puissance actuelle. Une étoile brillante, comme Capella, doit s'épuiser en une vingtaine de millions d'années, et quelques-unes des étoiles les plus brillantes (Rigel par exemple) ne peuvent pas durer plus d'un million d'années ou deux. Ainsi, les étoiles les plus brillantes sont forcément jeunes. De nouvelles étoiles sont peut-être en train de se former dans des régions de l'espace où la poussière est assez dense pour fournir la matière première.

En fait, en 1955, l'astronome américain George Herbig a détecté dans la poussière de la Nébuleuse d'Orion deux étoiles qui n'apparaissaient pas sur des photos de la même région prises quelques années plus tôt ; ces étoiles sont peut-être vraiment nées dans l'intervalle.

En 1965, on a repéré des centaines d'étoiles de température si faible qu'elles ne brillaient pas tout à fait. On les a détectées grâce à leur émission infrarouge, et on les appelle *géantes infrarouges,* parce qu'elles sont formées de grandes quantités de matière raréfiée. Vraisemblablement, ce sont des masses de poussières et de gaz en train de se rassembler, et qui deviennent de plus en plus chaudes. Elles finiront par devenir assez chaudes pour émettre de la lumière visible.

Les progrès suivants dans l'étude de l'évolution des étoiles sont venus de l'analyse des étoiles des amas globulaires. Dans un amas, les étoiles sont toutes à peu près à la même distance de nous, et par conséquent leur magnitude apparente est proportionnelle à leur magnitude absolue (comme c'est le cas pour les céphéides des Nuages de Magellan). Par conséquent, les magnitudes étant connues, on peut tracer un diagramme H-R de ces étoiles. On constate que les étoiles les moins chaudes (qui brûlent

leur oxygène lentement) se trouvent sur la séquence principale, mais que les plus chaudes ont tendance à s'en écarter. En raison de leur consommation rapide, et de leur vieillissement précoce, elles suivent une courbe montrant les stades successifs de leur évolution, d'abord vers les géantes rouges, puis vers la séquence principale à nouveau, à travers cette séquence, et enfin vers les naines blanches.

A partir de ces constatations et de considérations théoriques sur les combinaisons des particules subatomiques dans certaines conditions de température et de pression, Fred Hoyle a pu décrire en détail l'évolution d'une étoile. Selon lui, dans sa jeunesse, l'étoile varie peu en taille ou en température. C'est dans cette situation que se trouve – et continuera encore longtemps à se trouver – notre Soleil. Dans sa région centrale extrêmement chaude, l'étoile transforme son hydrogène en hélium, et l'hélium s'accumule en son centre. Quand ce noyau d'hélium atteint une certaine taille, l'étoile se met à changer de façon spectaculaire en taille et en température. Elle se dilate énormément, et sa température de surface diminue. Autrement dit, elle se déplace vers les géantes rouges en quittant la séquence principale. Plus l'étoile est massive, plus cela se produit tôt dans sa vie. Dans les amas globulaires, les étoiles les plus massives ont déjà parcouru des distances variables le long de cette route.

En dépit de sa température plus basse, la géante émet davantage de chaleur, à cause de l'augmentation de sa surface. Dans un lointain avenir, quand le Soleil quittera la séquence principale, et peut-être un peu avant, il sera devenu chaud au point que toute vie sera impossible sur la Terre. Cependant, cela ne se produira pas avant des milliards d'années.

Mais quel est précisément le changement dans le noyau d'hélium qui déclenche l'expansion de l'étoile en géante rouge ? Hoyle a suggéré que le noyau d'hélium lui-même se contractait, et par conséquent s'échauffait jusqu'à la température à laquelle les noyaux des atomes d'hélium fusionnent pour former du carbone, avec production d'une énergie supplémentaire. En 1959, le physicien américain David Elmer Alburger a montré en laboratoire que cette réaction était effectivement possible. C'est une réaction très rare et très peu probable, mais il y a tellement d'atomes d'hélium dans une géante rouge qu'un nombre suffisant de ces fusions peut se produire, libérant l'énergie nécessaire.

Hoyle va plus loin. Le jeune noyau central de carbone s'échauffe encore davantage, et des atomes encore plus compliqués, comme ceux de l'oxygène et du néon, commencent à se former. Pendant ce temps, l'étoile redevient plus petite et plus chaude, et revient vers la séquence principale. A ce stade, l'étoile a commencé à se structurer en couches successives, comme un oignon. Elle a un noyau d'oxygène et de néon, puis une couche de carbone, puis une couche d'hélium, et le tout est enveloppé dans une peau d'hydrogène qui n'est pas encore converti en autre chose.

Maintenant, après sa longue vie de consommatrice d'hydrogène, l'étoile arrive dans une période d'évolution rapide, une dégringolade à travers les combustibles qui lui restent. Sa vie ne peut plus être longue, car l'énergie produite par la fusion de l'hélium, et celles qui suivent, est faible par rapport à celle que fournit la fusion de l'hydrogène, environ vingt fois plus faible. En un temps relativement court, l'énergie vient à manquer, empêchant l'étoile de se contracter sous l'effet de son propre champ gravitationnel, et la contraction s'accélère. Non seulement l'étoile retrouve la taille d'une étoile normale, mais elle continue à se contracter, jusqu'à devenir une naine blanche.

Pendant la contraction, il peut arriver que les couches extérieures de l'étoile soient abandonnées, ou même expulsées dans l'espace à cause de la chaleur engendrée par la contraction. La naine blanche se retrouve ainsi entourée d'une « coquille » de gaz en expansion, visible dans nos télescopes sur ses bords, là où la quantité de gaz suivant la direction de visée est la plus grande. Ces naines blanches semblent entourées d'un petit « anneau de fumée » de gaz. On appelle ces anneaux *nébuleuses planétaires,* parce que l'anneau entoure l'étoile comme une orbite planétaire. L'anneau finit par s'étendre et se diluer jusqu'à disparaître, et on retrouve alors des naines blanches comme Sirius B, sans aucune nébulosité.

Les naines blanches se forment ainsi tranquillement, et c'est ce genre de mort paisible qui attend notre Soleil et les étoiles plus petites que lui. De plus, les naines blanches, si rien ne vient les déranger, ont devant elles une vie pratiquement indéfinie – une sorte de rigidité cadavérique persistante –, durant laquelle elles se refroidissent tout doucement, jusqu'à ne plus être assez chaudes pour émettre de la lumière (cela prend des milliards d'années) avant de passer des milliards et des milliards d'années dans la situation de naines noires.

D'autre part, si une naine blanche fait partie d'un système binaire, comme Sirius B ou Procyon B, et si l'autre étoile du système appartient à la séquence principale et se trouve très près de son compagnon, il peut y avoir des moments passionnants. A mesure que l'étoile de séquence principale se dilate, dans son évolution normale, une partie de la matière qui la constitue peut être attirée par le champ de gravitation intense de la naine, et s'installer en orbite autour d'elle. De temps en temps, une partie de cette matière atteint la surface de la naine, où le champ gravitationnel la comprime jusqu'à la fusion nucléaire. Si le morceau de matière qui est tombé sur la naine est assez gros, l'émission d'énergie peut être assez importante pour être visible de la Terre, et les astronomes détectent une nova... Naturellement, ceci peut se produire plusieurs fois de suite, et de fait on a enregistré des novae à répétition ou *novae récurrentes.*

Mais ce ne sont toujours pas des supernovae. Celles-ci, d'où viennent-elles ? Pour répondre à cette question, il faut nous tourner vers des étoiles nettement plus massives que notre Soleil. Elle sont relativement rares (dans toutes les catégories d'objets astronomiques, les gros sont plus rares que les petits). On estime qu'une étoile sur trente est plus massive que le Soleil. Mais même dans ce cas, il peut y avoir sept milliards d'étoiles de ce genre dans notre galaxie.

Dans les étoiles massives, le noyau est soumis à un champ gravitationnel plus grand que dans les étoiles ordinaires, et il est plus comprimé. Par conséquent il est plus chaud, et les réactions de fusion peuvent continuer au-delà de la limite oxygène-néon des étoiles plus petites. Le néon peut donner du magnésium, qui donne du silicium, et finalement du fer. A un âge avancé, l'étoile peut comporter plus d'une demi-douzaine de couches concentriques, dans chacune desquelles un combustible différent est consommé. La température au centre, à ce moment-là, peut avoir atteint trois ou quatre milliards de degrés. Une fois que l'étoile a commencé à former du fer, elle se trouve dans une impasse, parce que les atomes de fer représentent un point de stabilité maximum et d'énergie minimum. Pour transformer les atomes de fer en atomes plus compliqués, ou en atomes plus simples, il faut de toute façon apporter de l'énergie.

De plus, à mesure que la température centrale augmente, la pression de radiation augmente elle aussi, et comme la quatrième puissance de la température. Quand la température double, la pression de radiation est multipliée par seize, et son équilibre avec la gravitation devient de plus en plus délicat. Finalement, la température centrale peut devenir si élevée, d'après Hoyle, que les atomes de fer sont décomposés en hélium. Mais comme nous venons de le voir, il faut pour cela leur fournir de l'énergie ; la seule source possible est le champ gravitationnel de l'étoile. Quand celle-ci se ratatine, l'énergie gagnée peut convertir le fer en hélium. Mais l'énergie nécessaire est si grande que l'étoile doit se ratatiner prodigieusement, et atteindre une fraction minime de son volume initial ; selon Hoyle, cela doit se produire en une seconde environ.

Quand une étoile commence ainsi à se ratatiner, son noyau de fer est encore entouré d'un manteau volumineux d'atomes qui n'ont pas encore atteint la stabilité maximale. Quand les régions extérieures s'effondrent, et que leur température s'élève, ces substances qui sont encore combinables s'enflamment d'un seul coup. Le résultat est une explosion qui projette les matériaux externes loin du corps de l'étoile. Cette explosion est une supernova. C'est une explosion de ce type qui a créé la Nébuleuse du Crabe.

La matière projetée dans l'espace lors des supernovae a une importance énorme dans l'évolution de l'Univers. Au moment du big bang, il ne s'est formé que de l'hydrogène et de l'hélium. C'est dans le noyau des étoiles que se forment des atomes plus complexes – jusqu'aux atomes de fer. Sans les supernovae, ces atomes complexes resteraient dans les noyaux des étoiles, et finalement dans les naines blanches. Seule une quantité négligeable parviendrait, sous la forme des halos des nébuleuses planétaires, à se frayer un chemin vers l'espace.

Au cours des explosions qui donnent des supernovae, de la matière provenant des couches internes des étoiles est violemment projetée dans l'espace. L'énergie énorme dégagée par l'explosion permettrait même la formation d'atomes encore plus complexes que ceux du fer.

La matière projetée dans l'espace s'ajoute aux nuages de gaz et de poussières qui existent déjà, et sert de matière première pour fabriquer des *étoiles de seconde génération,* riches en fer et en métaux en général. Notre propre Soleil fait probablement partie de ces étoiles de seconde génération, beaucoup plus jeune que les vieilles étoiles des amas globulaires qui sont sans poussières. Ces *étoiles de première génération* sont pauvres en métaux et riches en hydrogène. La Terre, formée des mêmes débris que le Soleil, est extraordinairement riche en fer, un fer qui a peut-être existé au centre d'une étoile qui aurait explosé il y a des milliards et des milliards d'années...

Mais qu'arrive-t-il à la partie contractée des étoiles qui explosent en donnant des supernovae ? Forment-elles des naines blanches ? Étant plus grosses et plus massives que les étoiles ordinaires, forment-elles simplement des naines blanches plus grosses et plus massives ?

La première indication du contraire – nous ne pouvons pas nous attendre à des naines blanches de plus en plus massives – est venue en 1939 : l'astronome indien Subrahmanyan Chandrasekhar, travaillant à l'observatoire Yerkes près de Williams Bay dans le Wisconsin, a calculé qu'une étoile ayant plus de 1,4 fois la masse de notre Soleil (ce que l'on appelle depuis la *limite de Chandrasekhar*) ne peut pas devenir une naine blanche

par le processus « normal » décrit par Fred Hoyle. De fait, toutes les naines blanches observées jusqu'ici ont bien une masse inférieure à la limite de Chandrasekhar.

La raison de l'existence de cette limite est que les naines blanches ne peuvent pas se contracter davantage, à cause de la répulsion mutuelle des électrons (particules subatomiques dont nous parlerons au chapitre 7) contenus dans leurs atomes. Avec des étoiles plus massives, la gravitation est plus intense, et pour une masse 1,4 fois plus grande que celle du Soleil, la répulsion des électrons ne suffit plus, et la naine blanche s'effondre sur elle-même, pour former une étoile encore plus petite et encore plus dense, où les particules subatomiques sont pratiquement en contact. Pour détecter ces cas extrêmes, il a fallu de nouvelles méthodes d'exploration de l'Univers, utilisant des radiations différentes de la lumière visible.

Les fenêtres ouvertes sur l'Univers

Les armes principales dans la conquête du savoir sont la faculté de comprendre, et la curiosité insatiable qui la pousse toujours plus loin. Les ressources de l'esprit humain ont permis l'invention continuelle d'instruments qui ouvrent des horizons inaccessibles à nos seuls organes sensoriels.

LUNETTES ET TÉLESCOPES

L'exemple le plus connu est celui du « bond en avant » des connaissances à la suite de l'invention de la lunette en 1609. Lunettes et télescopes, avant tout, sont des « yeux géants ». La pupille de l'œil n'a que quelques millimètres de diamètre, tandis que le télescope du mont Palomar a cinq mètres de diamètre, c'est-à-dire une surface un million de fois plus grande environ : une étoile y apparaît un million de fois plus brillante qu'à l'œil nu. Ce télescope, entré en service en 1948, est le plus grand utilisé actuellement aux États-Unis, mais en 1976, l'Union Soviétique a mis au point un télescope de six mètres de diamètre, installé dans le Caucase.

C'est probablement la taille maximale plausible pour ce genre de télescope, et à vrai dire le miroir de six mètres n'est pas parfait. Mais il y a d'autres moyens d'améliorer les télescopes qu'avec des tailles de plus en plus grandes. Dans les années cinquante, Merle A. Ture a mis au point un tube électronique capable d'amplifier la lumière reçue par un télescope, multipliant sa puissance par trois. Des groupes de télescopes relativement petits, travaillant à l'unisson, donnent des images équivalentes à celles d'un télescope plus gros. Aussi bien aux États-Unis qu'en Union Soviétique, des systèmes de ce type sont en cours de réalisation, qui dépasseront de loin les télescopes géants de cinq et six mètres de diamètre. D'autre part, un gros télescope placé sur orbite autour de la Terre pourrait balayer le ciel sans être gêné par l'atmosphère, et donnerait de meilleurs résultats que n'importe quel télescope terrestre. Là aussi, la réalisation en est au stade des plans.

Mais grossissement et gain de luminosité ne sont pas les seuls apports du télescope à l'humanité. Le premier pas sur le chemin qui devait faire

du télescope plus qu'un entonnoir à lumière fut fait en 1666, quand Newton découvrit que la lumière pouvait être décomposée en donnant un *spectre* de couleurs. En faisant passer un rayon de soleil à travers un prisme triangulaire de verre, il s'aperçut que le pinceau lumineux s'étalait pour donner une bande irisée, où les couleurs passaient progressivement du rouge au violet, à travers le jaune, le vert et le bleu (figure 2.6). (Le phénomène était bien sûr connu depuis longtemps sous la forme de l'arc-en-ciel, produit par le passage des rayons du soleil à travers des gouttelettes qui agissent comme de petits prismes.)

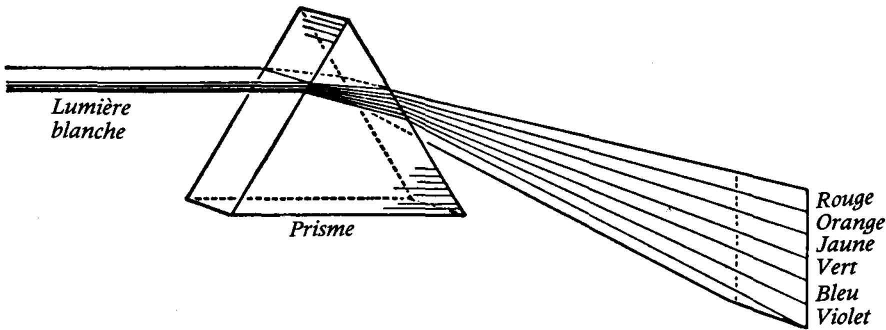


Figure 2.6. Expérience de Newton décomposant la lumière blanche.

Newton avait ainsi démontré que la lumière du Soleil, la *lumière blanche,* était un mélange de radiations particulières (identifiées depuis comme ayant des longueurs d'onde différentes) qui produisent sur l'œil l'effet de couleurs différentes. Le prisme sépare les couleurs, parce qu'en passant de l'air au verre et du verre à l'air, la lumière change de direction (*se réfracte*), et l'angle de réfraction dépend de la longueur d'onde – plus celle-ci est courte, plus la réfraction est importante. Les radiations violettes de courte longueur d'onde sont les plus réfractées, tandis que les rouges, qui correspondent aux plus grandes longueurs d'onde visibles, sont les moins réfractées.

Ce phénomène explique, entre autres choses, un défaut important des premières lunettes : les images qu'elles donnaient étaient entourées d'anneaux colorés qui en diminuaient la netteté. C'était un effet de la réfraction à travers le verre des lentilles.

Persuadé que c'était là un défaut inévitable des lunettes, Newton imagina de fabriquer un engin équivalent, mais sans lentille, où la lumière serait concentrée par un miroir parabolique. Les radiations de toutes les longueurs d'onde se réfléchissant de la même façon, on évitait de décomposer la lumière et de former des anneaux irisés. On échappait ainsi à ce défaut, nommé *aberration chromatique.*

En 1757, l'opticien anglais John Dollond imagina de fabriquer des lentilles de deux verres différents, de manière qu'elles donnent des irisations en sens inverse, susceptibles de se compenser. Il parvint ainsi à fabriquer des lentilles *achromatiques* (c'est-à-dire sans aberration chromatique). Grâce à cette invention, la lunette redevint populaire. La plus grande,

dont l'objectif a un diamètre d'un mètre, est à l'observatoire Yerkes, et sa construction date de 1897. On n'en a pas construit de plus grande depuis, et il est peu probable qu'on le fasse jamais, car des lentilles encore plus grandes absorberaient davantage de lumière que l'augmentation de leur diamètre ne permettrait d'en gagner. Les grands instruments actuels sont tous des télescopes : la surface de leurs miroirs absorbe très peu de lumière.

LE SPECTROSCOPE

En 1814, un opticien allemand, Joseph von Fraunhofer, imagina un perfectionnement du dispositif de Newton. Il fit passer le pinceau de lumière solaire par une fente étroite avant de l'envoyer sur le prisme. Le spectre obtenu était en réalité une série d'images de la fente données par les lumières de toutes les longueurs d'onde présentes dans la lumière solaire. Ces images étaient si nombreuses qu'elles se succédaient sans interruption et formaient un spectre continu, mais grâce à la perfection de son dispositif et à la finesse des images, Fraunhofer s'aperçut qu'il en manquait quelques-unes. Les raies noires dans le spectre correspondaient à des longueurs d'onde absentes dans le mélange solaire.

Fraunhofer nota l'emplacement des raies noires qu'il détectait, et en enregistra plus de sept cents. On les appelle depuis les *raies de Fraunhofer*. En 1842, les raies du spectre solaire furent photographiées pour la première fois par le physicien français Alexandre Edmond Becquerel. Ces photographies facilitaient considérablement l'étude. Avec les instruments modernes, on a détecté plus de trente mille raies dans le spectre solaire, et on a mesuré leurs longueurs d'onde.

En 1850, un certain nombre de savants commencèrent à se dire que ces raies étaient peut-être caractéristiques des différents éléments présents dans le Soleil. Les raies noires représenteraient l'absorption des radiations correspondantes par certains éléments. Vers 1859, les chimistes allemands Robert Wilhelm Bunsen et Gustav Robert Kirchhoff mirent au point une manière d'identifier ainsi les éléments. Ils chauffaient une substance jusqu'à l'incandescence, dispersaient en spectre la lumière émise, mesuraient la position des raies (qui dans ce cas étaient des raies lumineuses sur fond sombre), et faisaient correspondre chaque raie à un élément donné. Leur *spectroscope* devait rapidement permettre de découvrir de nouveaux éléments, donnant des raies absentes de tous les spectres des éléments connus. En un ou deux ans, Bunsen et Kirchhoff avaient ainsi découvert le césium et le rubidium.

En pointant le spectroscope vers le Soleil et les étoiles, on recueillit rapidement une masse étonnante de résultats chimiques et autres. En 1862, l'astronome suédois Anders Jonas Angström identifia les raies caractéristiques de l'hydrogène dans un spectre solaire.

On peut aussi détecter de l'hydrogène dans les autres étoiles, bien que les spectres des étoiles soient très différents les uns des autres, à cause, en particulier, des différences de composition chimique des étoiles. En fait, on peut classer les étoiles d'après la nature générale de leur système de raies spectrales. D'abord ébauchée en 1867 à partir de quatre mille spectres par l'astronome italien Pietro Angelo Secchi, cette classification devait être affinée vers 1890 par l'astronome américain Edward Charles Pickering, qui disposait de dizaines de milliers de spectres et de l'assistance minutieuse d'Annie J. Cannon et Antonia C. Maury.

A l'origine, la classification se faisait par lettres majuscules rangées dans l'ordre alphabétique, mais à mesure que l'on en apprenait davantage sur les étoiles, il devint nécessaire de modifier cela pour que l'ordre devienne logique. Si on range les étoiles par températures décroissantes, on trouve successivement les classes O, B, A, F, G, K, M, R, N et S. Chaque classe peut de plus être subdivisée au moyen de nombres de 1 à 10. Le Soleil, étoile de température moyenne, appartient à la classe G-O, tandis que l'étoile Alpha du Centaure est une G-2. Procyon, un peu plus chaude, est F-5, et Sirius, considérablement plus chaude, est A-O.

De même que le spectroscope permettait de déceler sur Terre de nouveaux éléments, il permettait d'en déceler dans le ciel. En 1868, l'astronome français Pierre Jules César Janssen, ayant observé en Inde une éclipse totale de Soleil, rapporta avoir remarqué une raie spectrale qui ne correspondait à aucun élément connu. L'astronome anglais Sir Norman Lockyer, sûr que cette raie indiquait la présence d'un nouvel élément, le baptisa *hélium,* du mot grec désignant le Soleil. C'est seulement trente ans plus tard que l'on devait découvrir de l'hélium sur la Terre.

Finalement, le spectroscope est devenu l'instrument de mesure de la vitesse radiale des étoiles, comme nous l'avons vu au début de ce chapitre, et le moyen d'étude de nombreux phénomènes – les propriétés magnétiques d'une étoile, sa température, la présence éventuelle d'un compagnon, etc.

De plus, les raies spectrales formaient une véritable encyclopédie de connaissances sur la structure atomique, mais on ne put s'en servir vraiment qu'à partir de 1890, après la découverte des particules subatomiques constituant l'atome. Par exemple, en 1885, le physicien allemand Johann Jakob Balmer montra que l'hydrogène produit toute une série de raies, espacées régulièrement suivant une formule simple. Une génération plus tard, ce résultat devait permettre des déductions importantes sur la structure de l'atome d'hydrogène (voir chapitre 8).

Lockyer lui-même montra que les raies spectrales produites par un élément donné se modifiaient à des températures élevées. Ceci révélait un changement dans les atomes. Encore une fois, on n'en profita pleinement que plus tard, après la découverte des particules encore plus petites qui forment l'atome, et dont certaines peuvent être chassées à haute température, ce qui change la structure de l'atome et les raies qu'il produit. (On a pris quelquefois ces raies modifiées pour des raies nouvelles, révélant la présence d'éléments nouveaux. Hélas, le seul que l'on ait découvert pour de bon dans le ciel est l'hélium.)

LA PHOTOGRAPHIE

Inventée en 1830 par l'artiste français Louis Jacques Mandé Daguerre, sous la forme des premiers *daguerréotypes,* la photographie ne devait pas tarder à se révéler un outil très précieux pour l'astronomie. Entre 1840 et 1850, plusieurs astronomes américains photographièrent la Lune, et l'une de ces photographies, due à l'astronome américain George Phillips Bond, fit sensation à la Grande Exposition de 1851 à Londres. On photographiait aussi le Soleil. En 1860, Secchi fit la première photographie d'une éclipse totale de Soleil. En 1870, et grâce aux photographies d'éclipses, il n'y avait plus de doute : la couronne et les protubérances appartenaient au Soleil, et pas à la Lune.

Pendant ce temps, et dès 1850, les astronomes photographiaient aussi les étoiles lointaines. En 1887, l'astronome écossais David Gill fit de la photographie stellaire une pratique courante. On pouvait déjà prévoir que la photographie deviendrait plus importante que l'observation directe dans l'exploration de l'Univers.

Peu à peu, la technique de la photographie par télescope s'améliorait. Le principal inconvénient était la petitesse du champ couvert par un télescope, dès que celui-ci dépasse une certaine taille. Si on essaie d'augmenter le champ, les bords de l'image se déforment. En 1930, l'opticien allemand d'origine russe Bernard Schmidt a inventé une lentille correctrice permettant d'éviter la distorsion au bord du champ. Avec un télescope équipé d'une telle lentille, on peut photographier à chaque fois un grand secteur du ciel, et y rechercher des objets intéressants, que l'on peut ensuite étudier de façon approfondie avec un télescope ordinaire. Puisque ces télescopes servent pratiquement toujours à faire des photos, on les appelle *chambres de Schmidt.*

La plus grande de ces chambres est un instrument de diamètre 1,35 mètre, mis en service en 1960 à Tautenberg, en Allemagne de l'Est. Ensuite vient l'instrument d'1,20 mètre utilisé à Palomar en liaison avec le télescope de cinq mètres. Le troisième, d'un mètre de diamètre, a été installé en 1961 dans un observatoire d'Arménie soviétique.

Vers 1800, William Herschel (l'astronome qui a deviné le premier la forme de notre galaxie) a réalisé une expérience très simple et très intéressante. Il a déplacé un thermomètre dans le spectre solaire obtenu avec un prisme, et constaté que le mercure s'élevait, même au-delà du rouge ! De toute évidence, il y avait des radiations invisibles, de longueur d'onde plus grande que celle du rouge. On leur donna le nom d'*infrarouges,* et nous savons maintenant qu'elles représentent 60 % du rayonnement solaire.

En 1801, le physicien allemand Johann Wilhelm Ritter explorait l'autre extrémité du spectre. Il s'aperçut que le nitrate d'argent, qui se décompose en donnant de l'argent métallique et en noircissant quand on l'expose à une lumière bleue ou violette, se décomposait encore plus vite si on le plaçait au-delà du violet du spectre ! C'est ainsi que Ritter découvrit la « lumière » que l'on appelle *ultraviolet.* A eux deux, Herschel et Ritter avaient élargi l'éventail de radiations connu depuis plus d'un siècle, et pénétré de nouveaux domaines de radiations.

Ces nouveaux domaines promettent une belle moisson d'informations. La partie ultraviolette du spectre, invisible pour nos yeux, se révèle avec une grande finesse à la photographie. En fait, avec un prisme de quartz (le quartz laisse passer l'ultraviolet, alors que le verre en absorbe la plus grande partie), on peut enregistrer un spectre ultraviolet très complexe, comme l'a fait pour la première fois en 1852 le physicien anglais George Gabriel Stokes. Malheureusement, l'atmosphère ne transmet que l'*ultraviolet proche* – celui dont la longueur d'onde est presque aussi grande que celle de la lumière violette. L'*ultraviolet lointain,* de longueur d'onde très courte, est absorbé dans la haute atmosphère.

LA RADIOASTRONOMIE

En 1860, le physicien écossais James Clerk Maxwell élabora une théorie prédisant une famille de radiations associées à des phénomènes électriques

et magnétiques (*radiations électromagnétiques*), famille dont la lumière visible ne constituait qu'une petite partie. La première confirmation expérimentale de cette prédiction n'est venue qu'un quart de siècle plus tard, sept ans après la mort prématurée de Maxwell des suites d'un cancer. En 1887, le physicien allemand Heinrich Rudolf Hertz, tirant un courant alternatif d'une bobine à induction, produisit et détecta une onde de longueur d'onde considérablement plus grande que celle de l'infrarouge ordinaire. On donna à ces ondes le nom de *radio.*

Les longueurs d'onde de la lumière visible peuvent se mesurer en *microns* (millioniémes de mètre). Elles vont de 0,39 micron (violet extrême) à 0,78 micron (rouge extrême). Ensuite vient le *proche infrarouge* (0,78 à 3 microns), l'*infrarouge moyen* (3 à 30 microns) et l'infrarouge lointain (30 à 1 000 microns). C'est ici que commencent les ondes radio : les *micro-ondes* vont de 1 000 à 160 000 microns, et les grandes ondes atteignent des kilomètres.

On peut aussi caractériser les radiations par leur *fréquence,* à savoir le nombre d'oscillations par seconde. La valeur obtenue est si grande pour la lumière visible et l'infrarouge qu'on utilise rarement la fréquence dans ces cas. Mais pour les ondes radio, la fréquence est moins grande et devient plus maniable. Mille oscillations par seconde donnent une fréquence d'un *kilocycle,* et un million d'oscillations par seconde un *mégacycle.* La région des micro-ondes va de 300 000 mégacycles à 1 000 mégacycles. Les ondes radio utilisées par les émetteurs normaux sont beaucoup plus longues, et leur fréquence est de l'ordre du kilocycle.

Moins de dix ans après la découverte de Hertz, l'autre bout du spectre se prolongeait d'une manière analogue. En 1895, le physicien allemand Wilhelm Konrad Roentgen découvrit accidentellement une radiation mystérieuse, à laquelle il donna le nom de *rayons X,* dont la longueur d'onde était plus courte que celle de l'ultraviolet. Plus tard, Rutherford mit en évidence des radiations de longueur d'onde encore plus faible, associées à la radioactivité, les *rayons gamma.*

La moitié du spectre correspondant aux courtes longueurs d'onde est maintenant subdivisée de la manière suivante : les longueurs d'onde de 0,39 à 0,17 micron correspondent à l'ultraviolet proche, celles de 0,17 à 0,01 micron à l'ultraviolet lointain, celles de 0,01 à 0,000 01 micron aux rayons X, et plus bas (jusqu'au milliardième de micron) elles correspondent aux rayons gamma.

Ainsi le spectre original de Newton s'étendait considérablement. Si nous appelons « octave » chaque doublement de la fréquence, comme on le fait pour le son, le spectre électromagnétique accessible à l'étude couvre presque soixante octaves. La lumière visible occupe une octave vers le milieu du spectre.

Bien entendu, ce spectre plus étendu enrichit notre vision des étoiles. Nous savons par exemple que la lumière solaire est riche en ultraviolet et en infrarouge. Notre atmosphère coupe la plus grande partie de ces radiations, mais en 1931, d'une façon tout à fait accidentelle, on découvrit une « fenêtre radio » sur l'Univers.

Karl Jansky, jeune ingénieur radio aux Laboratoires Bell, étudiait le bruit de fond qui accompagne toujours les transmissions radio. Il décela un bruit très faible et continu, qui ne pouvait venir d'aucune des sources connues. Il finit par décider que ce bruit était causé par des ondes radio venues de l'espace.

Tout d'abord, ce signal paraissait maximal dans la direction du Soleil. Mais de jour en jour, la direction du maximum et celle du Soleil s'écartaient l'une de l'autre. En 1933, Jansky put conclure que ces ondes venaient de la Voie lactée, et surtout de la région du Sagittaire, dans la direction du centre de la Galaxie.

C'est ainsi qu'est née la *radioastronomie.* Les astronomes ne s'y sont pas intéressés tout de suite, parce qu'elle avait plusieurs inconvénients sérieux. Elle ne donnait pas d'images nettes, seulement des pics sur une courbe, qui n'étaient pas faciles à interpréter. Et surtout, les ondes radio sont beaucoup trop longues pour permettre de distinguer une source aussi petite qu'une étoile. Les signaux radio avaient des longueurs d'onde des millions de fois plus grandes que celle de la lumière, et un récepteur ordinaire ne pouvait indiquer que la direction approximative des ondes. Un *radiotélescope* devrait avoir une antenne de réception un million de fois plus grande (en diamètre) que le miroir d'un télescope optique pour donner du ciel une image suffisamment précise. Une antenne équivalente au télescope de cinq mètres devrait avoir cinq mille kilomètres de diamètre, ce qui est manifestement impossible.

Ces difficultés masquèrent l'importance de la nouvelle découverte, mais un jeune radio amateur nommé Grote Reber s'y consacra, par pure curiosité. Il dépensa beaucoup d'argent, et passa toute l'année 1937 à construire dans sa cour un radiotélescope muni d'une antenne d'une dizaine de mètres de diamètre. Il commença ses observations en 1938, et découvrit un certain nombre de sources radio, en plus de celle du Sagittaire – une dans la constellation du Cygne, par exemple, et une autre dans Cassiopée. (On leur donna d'abord le nom d'*étoiles radio,* qu'elles fussent ou non des étoiles, mais on préfère maintenant les appeler *sources radio.*)

Pendant la Deuxième Guerre mondiale, les Anglais, qui mettaient au point le radar, s'aperçurent que le Soleil les gênait en émettant des micro-ondes. Ceci éveilla l'intérêt pour la radioastronomie, et après la guerre les Anglais continuèrent à étudier les émissions radio du Soleil. En 1950, ils avaient établi une corrélation entre ces signaux radio et les taches solaires. (Si Jansky avait détecté les émissions galactiques avant celles du Soleil, c'est parce qu'il avait fait ses expériences pendant une période d'activité solaire minimale.)

De plus, puisque la technique radar utilisait les mêmes longueurs d'onde que la radioastronomie, les astronomes disposaient à la fin de la guerre d'instruments adaptés à la manipulation des micro-ondes, qui n'existaient pas quelques années plus tôt. On les améliora rapidement, et l'intérêt suscité par la radioastronomie fit un grand bond en avant.

Les Anglais jouèrent le rôle de pionniers en construisant de grands radiotélescopes pour rendre la réception plus fine et distinguer des sources plus petites. C'est à Jodrell Bank, en Angleterre, que fut construite sous la direction de Sir Bernard Lovell une antenne de quatre-vingts mètres de diamètre, la première qui soit vraiment grande.

On a trouvé depuis des moyens d'affiner la réception sans devoir construire des antennes monstrueuses. On peut placer loin de l'autre deux télescopes de taille raisonnable. Si les deux antennes sont réglées avec des *horloges atomiques* extrêmement précises, et si elles se déplacent à l'unisson grâce à des commandes par ordinateur, l'ensemble des deux donne des résultats comparables à ceux que l'on obtiendrait avec un seul disque ayant

pour largeur la distance de séparation. De telles combinaisons d'antennes s'appellent des *antennes à longue base.* Les astronomes australiens, disposant de la place suffisante, ont été les premiers à réaliser de tels radiotélescopes, et maintenant des antennes travaillant en commun en Australie et en Californie donnent une base de plus de dix mille kilomètres.

Dans ces conditions, les radiotélescopes cessent d'être des « faiseurs de flou », ridicules par rapport à l'œil perçant des télescopes optiques. Ils peuvent en fait fournir plus de détails que les télescopes optiques. Bien sûr, on ne pourra guère établir sur terre de bases plus longues, mais les astronomes rêvent de radiotélescopes dans l'espace, coopérant entre eux et avec des antennes terrestres pour réaliser des bases encore plus longues.

D'ailleurs, bien avant de progresser jusqu'à leur niveau actuel, les radiotélescopes permettaient déjà des découvertes importantes. En 1947, l'astronome australien John C. Bolton parvint à préciser la position de la troisième source radio du ciel, ce qui permit de l'identifier : il s'agissait de la Nébuleuse du Crabe. De toutes les sources radio déjà repérées çà et là dans le ciel, c'était la première à être « accrochée » à un objet visible. Il semblait peu probable qu'une étoile donne naissance à une émission aussi intense, puisque les autres ne le faisaient pas. La source était beaucoup plus vraisemblablement le nuage de gaz en expansion dans la Nébuleuse.

Cette découverte vint renforcer d'autres indices tendant à montrer que les émissions radio cosmiques viennent surtout de masses de gaz turbulents. Le gaz turbulent de l'atmosphère externe du Soleil émet des ondes radio, ce qui rend le *Soleil radio* beaucoup plus gros que le Soleil visible. D'autre part, Jupiter, Saturne et Vénus, qui possèdent des atmosphères turbulentes, se sont aussi révélées être des sources d'émissions radio.

Jansky, qui était à l'origine de toutes ces découvertes, est resté pratiquement inconnu de son vivant (il est mort en 1950 à 44 ans, juste au moment où la radioastronomie démarrait pour de bon). Il connaît tout de même une gloire posthume : l'intensité des émissions radio cosmiques s'exprime en *janskys.*

VOIR AU-DELÀ DE NOTRE GALAXIE

La radioastronomie a permis de plonger très loin dans l'espace. Dans notre galaxie, il y a une puissante source radio, la plus puissante à l'extérieur du système solaire, située dans la constellation de Cassiopée. Walter Baade et Rudolph Minkowski ont pointé le télescope de cinq mètres de Palomar à l'emplacement indiqué pour cette source par les astronomes britanniques et y ont trouvé des traînées de gaz turbulent. Ce sont peut-être les débris de la supernova de 1572, que Tycho avait observée dans Cassiopée.

En 1951, on fit une découverte encore plus lointaine. La deuxième source radio (par ordre de puissance) se trouve dans la constellation du Cygne. Elle avait été signalée en 1944 par Reber. Quand les radiotélescopes ont commencé à la cerner plus précisément, on s'est aperçu que cette source était en dehors de notre galaxie – la première source localisée hors de la Voie lactée. Puis en 1951, observant cette région du ciel avec le télescope de Palomar, Baade y trouva une galaxie bizarre. Elle avait deux noyaux,

et paraissait tordue. Baade soupçonna aussitôt qu'il s'agissait d'une galaxie double, c'est-à-dire de deux galaxies, jointes par un bord comme deux cymbales. Il y vit deux galaxies en collision, une possibilité qu'il avait déjà discutée avec ses collègues. Son observation semblant corroborer cette idée, on accepta pendant quelque temps l'idée de galaxies en collision. La plupart des galaxies étant regroupées en amas, dans lesquels elles s'agitent comme des abeilles dans un essaim, de telles collisions ne semblaient pas invraisemblables.

On a pu localiser à 260 millions d'années-lumière de nous la source radio du Cygne, dont les signaux sont cependant plus forts que ceux de notre « voisine » du Crabe. C'était là la première indication que les radiotélescopes seraient capables de voir plus loin que les télescopes optiques. Même l'antenne de quatre-vingts mètres de Jodrell Bank, toute petite par rapport aux réalisations actuelles, dépassait déjà la portée du télescope de Palomar.

Mais, à mesure que le nombre de radio sources extragalactiques augmentait jusqu'à dépasser la centaine, le malaise des astronomes croissait. Ces sources ne pouvaient pas toutes correspondre à des collisions de galaxies.

En fait, l'idée même de collision de galaxies devenait moins plausible. L'astrophysicien soviétique Victor Amazaspovitch Ambartsumian avança en 1955, pour des raisons théoriques, l'idée que les galaxies explosaient plutôt que de se heurter.

Cette hypothèse fut considérablement renforcée par la découverte, en 1963, que la galaxie M 82, dans la constellation de la Grande Ourse (une source radio puissante située à dix millions d'années-lumière environ) était bien *une galaxie en explosion.* Le télescope de Palomar, utilisé à une certaine longueur d'onde, a révélé de grandes éruptions de matière (de mille années-lumière de long) sortant du centre de la galaxie. A partir de la quantité de matière expulsée, de la distance qu'elle a parcourue et de sa vitesse, l'explosion se serait produite il y a un million et demi d'années.

Il semble maintenant que les noyaux galactiques soient actifs, qu'il s'y passe des phénomènes turbulents d'une violence extrême, et que l'Univers en général soit beaucoup plus fantastique encore qu'on ne pouvait le penser avant l'avènement de la radioastronomie. La profonde sérénité du ciel, tel qu'il apparaît à l'œil nu, résulte des limitations de notre vue, qui ne nous montre que les étoiles les plus voisines, et pendant un temps limité.

Au centre même de notre galaxie, il y a une petite région, de quelques années-lumière tout au plus, qui est une source radio intense.

Par ailleurs, le fait que les galaxies en explosion existent, et que les noyaux galactiques actifs soient répandus (et peut-être universels) n'empêche pas forcément les collisions de galaxies de se produire. Dans chaque amas de galaxies, il semble vraisemblable que les grosses galaxies s'enrichissent aux dépens des petites, et il y a généralement dans l'amas une galaxie considérablement plus grosse que les autres. Certains indices inclinent à penser qu'elle a grossi en rencontrant et en absorbant des galaxies plus petites. On a photographié une grande galaxie possédant plusieurs noyaux, dont un seul lui appartenait à l'origine, les autres ayant autrefois fait partie de galaxies indépendantes. On commence à parler de *galaxies cannibales.*

Les nouveaux objets

En 1960, les astronomes pouvaient penser que les objets physiques de l'Univers ne leur réservaient plus beaucoup de surprises. Qu'il reste beaucoup à découvrir sur le plan théorique, de nouveaux points de vue, certainement. Mais trois siècles d'observations avec des instruments de plus en plus perfectionnés ne laissaient pas espérer la découverte de types nouveaux de galaxies, d'étoiles ou de quoi que ce soit d'autre.

Pour ceux qui pensaient ainsi, le choc a dû être rude. Tout a commencé par les recherches menées au sujet de certaines sources radio qui semblaient un peu hors de l'ordinaire, sans plus.

LES QUASARS

Les premières sources radio lointaines semblaient être rattachées à des nuages étendus de gaz turbulent : la Nébuleuse du Crabe, des galaxies lointaines, etc. Mais quelques sources semblaient anormalement petites. A mesure que les radiotélescopes se perfectionnaient et donnaient des images plus fines, on commençait à envisager la possibilité que ces sources soient des étoiles.

Parmi ces sources compactes, quelques-unes étaient répertoriées dans le troisième catalogue de Cambridge, réalisé par l'astronome anglais Martin Ryle et ses collaborateurs, sous les numéros 3C48, 3C147, 3C196, 3C273 et 3C286 (3C pour « troisième » et « Cambridge »).

En 1960, l'astronome américain Allen Sandage se mit à passer au peigne fin, avec le télescope de cinq mètres, les régions contenant ces sources, et dans chaque cas il trouva une étoile qui semblait effectivement être la source radio. La première correspondait à 3C48. Quant à 3C273, le plus brillant de ces objets, sa position précise put être déterminée par Cyril Hazard en Australie, qui enregistra son « extinction radio » quand la Lune passa devant lui.

Les étoiles correspondantes avaient été enregistrées sur les atlas célestes établis auparavant, et on n'y avait vu que des membres faibles de notre galaxie. Mais des photographies minutieuses, résultant de leur émission radio inhabituelle, révélèrent bientôt que les choses n'étaient pas si simples. On détecta de faibles nébulosités autour de certains de ces objets, et même un jet de matière sortant de 3C273. En fait, 3C273 comportait deux sources radio, l'une liée à l'étoile et l'autre à ce jet de matière. D'autre part, une étude soigneuse révéla que ces étoiles étaient anormalement riches en ultraviolet.

On se mit donc à penser que ces sources radio compactes, bien que ressemblant à des étoiles, pouvaient ne pas être des étoiles ordinaires. On les baptisa *sources radio quasi stellaires* (« quasi-stellar radio sources » en anglais), nom abrégé dès 1964 en *quasars* par le physicien américain d'origine chinoise Hong Yee Chiu. Malgré sa sonorité dure, le terme de « quasar » est maintenant bien établi dans le vocabulaire astronomique.

De toute évidence, les quasars étaient assez intéressants pour mériter qu'on mobilise à leur sujet tout l'arsenal des techniques astronomiques, y compris la spectroscopie. Pour obtenir leurs spectres, il fallut la collaboration de nombreux astronomes, en particulier Allen Sandage, Jesse

L. Greenstein et Maarten Schmidt. Quand ils obtinrent les spectres, en 1960, ils se trouvèrent devant des raies bizarres, qu'ils ne parvenaient pas à identifier. De plus, les raies du spectre d'un quasar donné n'avaient rien de commun avec celles d'un autre.

En 1963, Schmidt revint au spectre de 3C273, qui, étant le plus brillant des quasars, donnait le spectre le plus net. Il y repéra six raies, dont quatre étaient espacées de telle manière qu'elles ressemblaient à une série de raies de l'hydrogène, mais anormalement situées dans le spectre. Et si ce déplacement était dû à un décalage vers le rouge ? Si c'était le cas, il s'agissait d'un décalage énorme, indiquant une vitesse d'éloignement de plus de 40 000 km/s. Ceci semblait incroyable, et pourtant, on pouvait ainsi identifier les deux autres raies : l'une produite par l'oxygène privé de deux électrons, l'autre par le magnésium privé de deux électrons.

Schmidt et Greenstein se tournèrent alors vers les spectres des autres quasars, et constatèrent que là encore on pouvait identifier les raies, à condition d'admettre d'énormes décalages vers le rouge.

Ces décalages pouvaient résulter de l'expansion de l'Univers. Mais en appliquant la loi de Hubble au calcul des distances, il se révélait que les quasars ne pouvaient pas du tout appartenir à notre galaxie. C'était au contraire les objets les plus lointains de l'Univers, situés à des milliards d'années-lumière.

A la fin des années soixante, des recherches systématiques avaient identifié cent cinquante quasars, et on avait pu étudier les spectres de cent dix d'entre eux. Chacun d'eux, sans exception, présentait un énorme décalage vers le rouge, plus énorme encore que celui de 3C273. Quelques-uns d'entre eux doivent se trouver à neuf milliards d'années-lumière.

Si les quasars sont effectivement aussi éloignés que leur décalage vers le rouge le laisse supposer, ils confrontent les astronomes à un certain nombre d'énigmes. D'abord, ils doivent être prodigieusement lumineux pour apparaître aussi brillants à de telles distances, de trente à cent fois plus lumineux qu'une galaxie ordinaire tout entière.

Mais si c'est le cas, et si les quasars ont la forme des galaxies, ils devraient contenir jusqu'à cent fois plus d'étoiles qu'une galaxie ordinaire, et être cinq ou six fois plus grands dans chacune de leurs dimensions. Même à la distance énorme où ils se trouvent, les grands télescopes devraient les montrer sous la forme de taches étendues. Ce n'est pas le cas. Dans le plus gros télescope, ils apparaissent comme des points, et par conséquent, malgré leur luminosité exceptionnelle, ils pourraient bien être beaucoup plus petits que les galaxies ordinaires.

Cette petitesse est encore accentuée par un autre phénomène. Dès 1963, on s'était aperçu que l'énergie émise par les quasars était variable, tant dans le domaine de la lumière que dans celui des ondes radio. Des augmentations et diminutions de l'ordre de trois magnitudes se produisaient en quelques années.

Pour que son émission varie autant en si peu de temps, un objet doit être petit. Des petites variations pourraient provenir de phénomènes limités à une petite région de l'objet, mais des variations aussi importantes doivent l'impliquer dans son ensemble. Si c'est le cas, quelque chose doit se faire sentir à travers tout l'objet en un temps inférieur à celui de la variation. Mais aucun effet ne peut voyager plus vite que la lumière. Un quasar qui varie sensiblement en quelques années ne peut pas avoir un

diamètre de plus de quelques années-lumière. En fait, certains calculs donnent aux quasars un diamètre de l'ordre de la semaine-lumière seulement (même pas mille milliards de kilomètres !).

Des corps aussi petits et aussi lumineux doivent dépenser de l'énergie à une cadence telle que leurs réserves ne peuvent pas durer longtemps (à moins, ce qui est possible, qu'il existe une source d'énergie que nous ne soyons pas encore capables d'imaginer). Certains calculs permettent d'estimer qu'un quasar ne peut émettre à cette cadence que pendant un million d'années environ. Dans ce cas, les quasars que nous voyons ne sont devenus quasars que très récemment, à l'échelle cosmique. Et il doit exister des objets qui ont été des quasars et ont cessé de l'être.

En 1965, Sandage annonça la découverte d'objets qui pouvaient bien être d'anciens quasars. Ils ressemblaient à des étoiles ordinaires bleuâtres, mais possédaient d'énormes décalages vers le rouge. Ils étaient aussi lointains, aussi petits et aussi lumineux que des quasars, mais ils n'émettaient pas d'ondes radio. Sandage les appela *objets stellaires bleus* ou *OSB.*

Les OSB semblent plus nombreux que les quasars : on a pu estimer en 1967 que 100 000 d'entre eux devaient se trouver à portée de nos télescopes. Il existe beaucoup plus de ces objets que de quasars, parce qu'ils durent beaucoup plus longtemps sous cette forme.

Les astronomes ne sont pas unanimes à estimer les quasars très lointains. Leurs énormes décalages vers le rouge pourraient ne pas être *cosmologiques,* c'est-à-dire ne pas être liés à l'expansion de l'Univers. Ce pourraient être des objets relativement proches, s'éloignant de nous pour une raison inconnue, locale – par exemple après avoir été éjectés d'un noyau galactique à une vitesse fantastique.

Le partisan le plus enthousiaste de cette idée est l'astronome américain Halton C. Arp, qui a montré des connections entre certains quasars et des galaxies proches : celles-ci ayant des décalages vers le rouge assez faibles, le décalage considérable vers le rouge des quasars (nécessairement voisins des galaxies s'il y a vraiment une relation entre eux) doit avoir une autre origine que l'éloignement.

Une autre énigme a surgi à la fin des années soixante-dix, lorsqu'on découvrit que des sources radio distinctes à l'intérieur d'un quasar (que les radiotélescopes à longue base permettent de distinguer) semblaient s'éloigner les unes des autres à des vitesses plusieurs fois supérieures à celle de la lumière. Ceci étant impossible, dans l'état actuel de la physique, il faudrait donc considérer les quasars comme beaucoup plus proches, ce qui ramènerait les vitesses observées au-dessous de la vitesse de la lumière.

Cependant, l'idée que les quasars pourraient être relativement proches (et donc moins lumineux et prodigues en énergie, ce qui résoudrait aussi cette énigme-là) ne fait pas l'unanimité parmi les astronomes. On pense en général que le décalage vers le rouge est bien d'origine cosmologique, que les relations quasars-galaxies décelées par Arp ne sont pas fondées, et que l'observation des vitesses supérieures à celle de la lumière est une illusion d'optique (pour laquelle on a déjà présenté plusieurs explications plausibles).

Si donc les quasars sont aussi éloignés que le laisse supposer leur décalage vers le rouge, s'ils sont aussi petits et lumineux que ces distances le laissent croire, qu'est-ce qu'un quasar ?

La réponse la plus vraisemblable remonte à 1943, quand l'astronome américain Carl Seyfert a repéré une galaxie bizarre, avec un noyau très petit et très brillant. On en a découvert d'autres depuis, et on les appelle *galaxies de Seyfert.* Bien qu'on en ait seulement dénombré une douzaine à la fin des années soixante, certaines raisons permettent de supposer qu'une galaxie sur cent pourrait être une galaxie de Seyfert.

Ces galaxies pourraient-elles être intermédiaires entre les galaxies ordinaires et les quasars ? Leur centre brillant présente des variations qui leur donnent une taille de l'ordre de celle des quasars. Si ce centre était encore renforcé, et si le reste de la galaxie était encore affaibli, on ne les distinguerait plus des quasars. L'une des galaxies de Seyfert, 3C120, a presque l'aspect d'un quasar.

Les galaxies de Seyfert n'ont que de faibles décalages vers le rouge, et ne sont pas très éloignées. Se pourrait-il que les quasars soient des galaxies de Seyfert si éloignées que nous ne pourrions voir que leur noyau, petit et lumineux ? Et qu'ils soient si éloignés que nous ne verrions que les galaxies les plus grandes, nous donnant l'impression que les quasars sont extraordinairement lumineux ?

De fait, des photographies récentes ont révélé des traces de nébulosités autour des quasars, semblant suggérer la galaxie, faible et ténue, qui entoure le centre petit, actif et très lumineux. On peut donc supposer que les régions lointaines de l'Univers, au-delà d'un milliard d'années-lumière, sont aussi riches en galaxies que les régions proches. Mais la plupart de ces galaxies seraient trop faibles pour être visibles, et nous ne verrions que les centres brillants des plus actives et des plus grandes d'entre elles.

LES ÉTOILES À NEUTRONS

Tandis que les ondes radio avaient révélé l'énigme des quasars, l'autre bout du spectre allait nous conduire à un autre type d'objet, tout aussi bizarre.

En 1958, l'astrophysicien américain Herbert Friedman découvrit que le Soleil émettait une quantité considérable de rayons X. On ne pouvait pas les détecter depuis la Terre, parce que l'atmosphère les absorbe. Mais on pouvait facilement les détecter avec des fusées emmenant les instruments appropriés hors de l'atmosphère.

Pendant quelque temps, l'origine des rayons X solaires est restée mystérieuse. La température de la surface du Soleil n'est que de 6 000° C – assez pour vaporiser n'importe quelle matière mais pas assez pour produire des rayons X. La source devait se trouver dans la couronne du Soleil, le halo ténu de gaz qui entoure le Soleil sur des millions de kilomètres. Bien que la couronne émette une lumière équivalente à la moitié de la pleine lune, elle est complètement masquée par la lumière du Soleil lui-même, et n'est visible que pendant les éclipses, tout au moins dans des conditions normales. Mais en 1930, l'astronome français Bernard Ferdinand Lyot inventa un instrument permettant, au moins en altitude et par beau temps, d'observer la couronne interne même en l'absence d'éclipse.

On y plaçait la source des rayons X, car même avant les observations faites hors de l'atmosphère, on soupçonnait la couronne d'être le siège de températures anormalement élevées. Des études du spectre de la couronne

(enregistré lors des éclipses) avaient révélé des lignes n'appartenant à aucun élément connu. On avait donc imaginé l'existence d'un nouvel élément, que l'on avait baptisé *coronium*. Mais en 1941, on avait montré que les raies du coronium pouvaient être produites par des atomes de fer ayant perdu un certain nombre de particules subatomiques. Pour les leur enlever, il fallait une température de l'ordre du million de degrés, certainement suffisante pour produire des rayons X.

L'émission de rayons X augmente brusquement quand une protubérance solaire fait irruption dans la couronne. L'intensité de l'émission dans ce cas implique des températures de l'ordre de cent millions de degrés dans la partie de la couronne qui surmonte la protubérance. La raison de l'apparition de températures aussi élevées dans le gaz ténu qui constitue la couronne est encore controversée. (Le mot « température » doit être pris ici dans un sens différent du sens habituel. Il s'agit d'une mesure de l'énergie cinétique des atomes dans le gaz, mais comme ceux-ci sont peu nombreux, l'énergie totale par unité de volume est relativement faible. Les rayons X sont produits par les collisions entre particules d'énergie très élevée.)

Il y a des rayons X qui viennent aussi de régions situées plus loin que le Soleil. En 1963, des instruments ont été emmenés par une fusée, à l'initiative de Bruno Rossi et d'autres astronomes, pour voir si la surface de la Lune réfléchissait les rayons X émis par le Soleil. Au lieu de cela, ces instruments ont détecté deux sources particulièrement concentrées de rayons X dans d'autres régions du ciel. La plus faible (Tau X-1, parce qu'elle est située dans la constellation du Taureau) devait rapidement être identifiée : il s'agissait encore une fois de la Nébuleuse du Crabe. En 1966, on pouvait associer la deuxième, plus forte, dans la constellation du Scorpion (Sco X-1), à un objet optiquement visible qui semblait être (comme la Nébuleuse du Crabe) le reste d'une vieille nova. Depuis lors, on a détecté dans le ciel de nombreuses sources de rayons X.

Pour émettre des rayons X de haute énergie, avec une intensité suffisante pour qu'ils soient détectables à des distances interstellaires, il fallait une source de grande masse et de température extrêmement élevée. Une émission analogue à celle de la couronne solaire aurait été tout à fait insuffisante.

Une masse importante, jointe à une température d'un million de degrés, cela suggère un objet encore plus condensé qu'une naine blanche. Dès 1934, Zwicky avait suggéré que les particules subatomiques d'une naine blanche pouvaient, dans certaines conditions, se combiner pour donner des particules sans charge électrique appelées *neutrons*. Ceux-ci pouvaient être pressés les uns contre les autres jusqu'à être effectivement en contact. Le résultat serait une sphère de vingt kilomètres de diamètre, possédant la masse d'une étoile. En 1939, le physicien américain J. Robert Oppenheimer calcula de façon détaillée les propriétés que devrait avoir une telle étoile. Elle atteindrait, au moins pendant les premières phases de son existence, une température assez élevée pour émettre des rayons X en quantité.

Friedmann a limité sa recherche d'une telle étoile à la région de la Nébuleuse du Crabe, où l'on pensait que l'explosion fantastique responsable de l'existence de la nébuleuse pourrait bien avoir laissé derrière elle, au lieu d'une naine blanche condensée, une étoile à neutrons superdense. En juillet 1964, la Lune devait passer devant la Nébuleuse du Crabe, et

on lança une fusée pour enregistrer l'émission X de celle-ci. Si cette émission venait d'une étoile à neutrons, elle serait éteinte d'un seul coup par l'interposition de la Lune, tandis que si la nébuleuse tout entière émettait, l'extinction serait graduelle. C'est ce qui s'est produit, et on en a conclu que la nébuleuse se comportait comme la couronne solaire – en plus grand et en plus puissant.

Pendant quelque temps, l'existence des étoiles à neutrons et la possibilité de les détecter ont semblé moins probables. Mais la même année où la Nébuleuse du Crabe s'était révélée décevante, une découverte nouvelle est venue d'une autre direction. Les ondes radio émises par certaines sources ont montré de très rapides fluctuations en intensité. C'était comme s'il y avait des « clignotants radio » çà et là.

Très vite, les astronomes ont mis au point des instruments capables de détecter des impulsions très courtes, susceptibles de permettre des études plus détaillées de ces changements rapides. L'un des astronomes à utiliser un tel radiotélescope était Anthony Hewish, à l'observatoire de l'université de Cambridge. Il a dirigé la construction de 2 048 détecteurs indépendants, couvrant une surface de près de deux hectares, et en juillet 1967 le dispositif a été mis en service. En moins d'un mois, une jeune assistante britannique, Jocelyn Bell, a détecté des impulsions radio provenant d'un point à mi-chemin entre Véga et Altaïr. Cette source n'était pas difficile à détecter, et l'aurait été plus tôt si les astronomes s'étaient attendus à trouver des impulsions aussi courtes, et avaient mis au point l'équipement nécessaire à leur détection. Dans ce cas, les impulsions étaient étonnamment courtes, et ne duraient qu'un trentième de seconde. Chose plus étonnante encore, elles se succédaient avec une régularité remarquable à des intervalles de 1,33 seconde. Les intervalles, en fait, étaient si réguliers que l'on a pu déterminer leur valeur à 1/100 000 000 de seconde près : 1,33730109 seconde.

Bien entendu, il n'y avait pas moyen, au départ, de dire ce que représentaient ces impulsions. Hewish n'a pu qu'inventer le nom d'*étoile pulsante* (« pulsating star ») rapidement abrégé en *pulsar,* nom sous lequel ce nouvel objet est désormais désigné.

Il faudrait d'ailleurs dire « nouveaux objets » au pluriel, car après que J. Bell eut trouvé le premier, Hewish en chercha d'autres. Quand il annonça sa découverte, en février 1968, il en avait découvert quatre, ce qui lui a valu une part du prix Nobel de physique en 1974. D'autres astronomes se mirent en chasse, et on connaît maintenant quatre cents pulsars. Il est possible qu'il en existe dans notre galaxie plus de cent mille. Certains se trouvent peut-être à une centaine d'années-lumière. (Il n'y a aucune raison de supposer que les autres galaxies en soient dénuées, mais la distance est trop grande pour que nous puissions les détecter.)

Tous les pulsars sont caractérisés par une extrême régularité de leurs pulsations, mais leur période varie d'un pulsar à l'autre. L'un d'eux a une période de 3,7 secondes, tandis qu'un autre (repéré dans la Nébuleuse du Crabe en novembre 1968 par les astronomes de Green Bank, Virginie) émet trente impulsions par seconde : sa période vaut 0,033089 seconde.

Une question évidente se pose immédiatement : qu'est-ce qui peut produire des oscillations aussi rapides avec une telle régularité ? Il faut qu'un corps astronomique subisse des changements à des intervalles aussi courts que ceux des impulsions. Cela pourrait-il être une planète, tournant

autour d'une étoile et émettant une puissante impulsion radio quand elle émerge, à chaque révolution ? Ou une planète tournant sur elle-même ne pourrait-elle pas avoir à sa surface une région émettant une quantité considérable d'ondes radio, région qui passerait rapidement devant nous à chaque rotation de la planète ?

Seulement, cette planète devrait tourner – sur elle-même ou autour de son étoile – en un temps de l'ordre de la seconde ou de la fraction de seconde, ce qui est impensable. Pour que les impulsions soient aussi rapides que celles des pulsars, il faut qu'un objet tourne – sur lui-même ou autour d'un autre – à une vitesse fantastique, ce qui exige une très petite taille, combinée à des températures très élevées, ou un champ gravitationnel énorme, ou les deux à la fois.

Ceci évoque immédiatement les naines blanches, mais même elles ne sauraient tourner assez vite sur elles-mêmes, ou l'une autour de l'autre, ou produire d'une manière ou d'une autre des impulsions, à une fréquence aussi grande. Les naines blanches sont trop grandes et ont un champ gravitationnel trop faible pour permettre de rendre compte des pulsars.

Thomas Gold se rendit compte immédiatement qu'il y avait dans cette affaire une étoile à neutrons. Elle est assez petite et dense pour tourner sur elle-même en moins de quelques secondes. De plus, la théorie avait déjà prévu que les étoiles à neutrons devaient avoir des champs magnétiques énormes, avec des pôles qui n'étaient pas nécessairement situés aux pôles de rotation. Les électrons seraient si fortement liés par ces champs magnétiques qu'ils ne pourraient émerger qu'aux pôles magnétiques. Lors de leur éjection, ils perdraient de l'énergie sous forme d'ondes radio. Il y aurait donc un puissant « émetteur » radio à chacun des deux pôles magnétiques de l'étoile à neutrons.

Si, au cours de la rotation de l'étoile, l'un de ces faisceaux d'ondes (ou les deux) balaie notre direction, nous détectons une impulsion très courte à son passage, une fois (ou deux fois) par tour. S'il en est ainsi, nous ne détectons que les pulsars dont l'une des directions d'émission se trouve balayer la Terre à chaque tour. Certains astronomes estiment qu'une étoile à neutrons sur cent remplirait cette condition. S'il y a vraiment cent mille étoiles à neutrons dans notre galaxie, nous ne pouvons en détecter que mille.

Gold est allé plus loin. Il a fait remarquer que, si sa théorie était correcte, l'étoile à neutrons perdrait de l'énergie en rayonnant ces ondes, et que sa vitesse de rotation diminuerait. Suivant cette idée, plus la période d'un pulsar est courte, plus il est jeune, plus il perd rapidement de l'énergie, et plus son ralentissement est important.

Le pulsar le plus rapide connu à l'époque était dans la Nébuleuse du Crabe. Ce pouvait fort bien être le plus jeune, puisque l'explosion de la supernova, soupçonnée d'avoir laissé derrière elle une étoile à neutrons, s'était produite moins de mille ans auparavant.

On a étudié avec soin la période de ce pulsar, et on a effectivement constaté qu'elle augmentait, comme Gold l'avait prédit. La période augmentait de 36,48 milliardièmes de seconde chaque jour. On a découvert le même phénomène pour d'autres pulsars, et au début des années soixante-dix l'hypothèse de l'étoile à neutrons était largement admise.

Il arrive parfois qu'un pulsar accélère brusquement un peu, puis retrouve son rythme lent normal. Certains astronomes voient là le résultat d'un *tremblement d'étoile*, le rééquilibrage des masses dans le corps de l'étoile. Ce pourrait aussi être le résultat de l'arrivée sur l'étoile d'un corps assez massif pour faire varier sensiblement sa vitesse de rotation.

Il n'y a aucune raison que toute l'énergie perdue par les électrons le soit sous forme d'ondes radio. Le phénomène devait produire une émission à toutes les longueurs d'onde du spectre, et en particulier de la lumière visible.

L'attention s'est concentrée sur la région de la Nébuleuse du Crabe, où on pouvait espérer voir des restes de l'explosion. Et de fait, en janvier 1969, on a constaté qu'une étoile faible située dans la nébuleuse clignotait exactement au même rythme que le pulsar. On l'aurait détectée plus tôt si les astronomes avaient pu penser qu'il serait intéressant de chercher des clignotements aussi rapides. Le pulsar de la Nébuleuse du Crabe a donc été le premier pulsar optique découvert – la première étoile à neutrons visible.

Ce pulsar émettait aussi des rayons X. 5 % environ de tous les rayons X émis par la nébuleuse tout entière venaient de ce point minuscule. C'est ainsi que la relation entre rayons X et étoiles à neutrons, qui semblait presque abandonnée en 1964, a connu une résurrection triomphale.

On aurait pu croire que les étoiles à neutrons ne réservaient plus aucune surprise. Mais en 1982, des astronomes utilisant le radiotélescope de trois cents mètres de diamètre situé à Arecibo (Porto Rico) ont découvert un pulsar qui oscillait 642 fois par seconde, vingt fois plus vite que celui de la Nébuleuse du Crabe. Il est probablement plus petit que la plupart des pulsars, et n'a sans doute qu'un diamètre de l'ordre de cinq kilomètres, avec une masse deux ou trois fois supérieure à celle de notre Soleil. On imagine l'intensité de son champ gravitationnel. Même ainsi, une rotation aussi rapide ne doit pas être loin de le faire voler en éclats. Autre énigme : sa rotation diminue beaucoup moins vite que ne le laisse prévoir la cadence à laquelle il émet de l'énergie.

On a détecté depuis un autre *pulsar rapide*, et les astronomes s'interrogent sur ce qui a pu lui donner naissance.

LES TROUS NOIRS

Même l'étoile à neutrons ne représente pas la plus grande densité concevable. Quand Oppenheimer a calculé en 1939 les propriétés que devrait avoir une étoile à neutrons, il a signalé également qu'une étoile assez massive (plus de 3,2 fois la masse de notre Soleil) pouvait s'écraser sur elle-même jusqu'à devenir un point, une *singularité*.

Quand l'effondrement dépasse le stade de l'étoile à neutrons, le champ gravitationnel devient si intense, suivant cette théorie, qu'aucune matière, ni aucune radiation, ne peut s'en échapper. Ainsi, n'importe quel objet piégé dans le champ gravitationnel inimaginable d'un tel monstre y tomberait sans espoir de retour, ce qui évoque l'image d'un « trou » infiniment profond dans l'espace. Et puisque même la lumière ne pourrait pas s'en échapper, ce serait un *trou noir*, terme utilisé pour la première fois dans les années soixante par le physicien américain John Archibald Wheeler.

Une étoile sur mille seulement est assez massive pour avoir une chance, en s'effondrant sur elle-même, de former un trou noir. Et parmi ces étoiles, la plupart doivent perdre assez de masse, au cours de l'explosion en supernova, pour échapper à ce sort. Mais même dans ces conditions, il peut y avoir des dizaines de millions d'étoiles de cette catégorie dans notre galaxie, et au cours de son existence, il a pu y en avoir des milliards. Si une sur mille seulement de ces étoiles lourdes finit par former un trou noir, il devrait y avoir un million de ceux-ci d'un bout à l'autre de la Galaxie. Si c'est le cas, où sont-ils ?

Le problème est qu'ils sont extrêmement difficiles à détecter. On ne peut pas les voir de la manière habituelle, puisqu'ils ne peuvent émettre ni lumière ni aucune forme de radiations. Et si leur champ gravitationnel est prodigieusement grand à leur voisinage immédiat, il n'est pas plus grand que celui d'une étoile ordinaire à des distances stellaires.

Dans certains cas, pourtant, un trou noir pourrait se trouver dans des conditions rendant la détection possible. Supposons qu'un trou noir fasse partie d'une « étoile double », c'est-à-dire que son compagnon et lui tournent autour de leur centre de masse commun, le compagnon étant une étoile normale.

Si les deux compagnons sont assez proches l'un de l'autre, de la matière peut quitter peu à peu l'étoile normale, et s'installer en orbite autour du trou noir, formant ce que l'on appelle un *disque d'accrétion*. Peu à peu, la matière contenue dans ce disque descendrait en spirale vers le trou noir, et ce faisant, par un processus bien connu, elle émettrait des rayons X.

Il faut donc chercher dans le ciel une source X invisible, semblant tourner autour d'une autre étoile, visible celle-ci.

En 1965, on a détecté une source X particulièrement intense dans la constellation du Cygne, à laquelle on a donné le nom de Cygne X-1. On évalue sa distance à dix mille années-lumière. C'était une source X comme les autres, jusqu'à ce qu'en 1970 un satellite soit lancé du Kenya pour détecter des sources X (il en a découvert 161). En 1971, ce satellite a permis de détecter des changements irréguliers dans l'émission de Cygne X-1, changements comme ceux que produirait l'arrivée dans un trou noir de « bouffées » de matière du disque d'accrétion.

On se mit aussitôt à étudier avec le plus grand soin Cygne X-1, et on trouva dans son voisinage immédiat une grosse étoile bleue, très chaude, environ trente fois plus massive que notre Soleil. L'astronome C.T. Bolt, de l'université de Toronto, a pu montrer que cette étoile et Cygne X-1 tournaient l'un autour de l'autre. A partir des éléments de son orbite, Cygne X-1 devait être cinq à huit fois plus massif que notre Soleil. Si c'était une étoile normale, on l'aurait vue. Puisqu'on ne la voyait pas, ce devait être un très petit objet. Trop massif pour être une naine blanche, ou même une étoile à neutrons, ce devait être un trou noir. Les astronomes ne sont pas encore tout à fait sûrs de ce résultat, mais la plupart d'entre eux pensent bien que Cygne X-1 est le premier trou noir effectivement découvert.

Les trous noirs, semble-t-il, pourraient se former de préférence aux endroits où les étoiles sont le plus serrées, et où de grandes masses de matière ont le plus de chances de s'accumuler. Or, les régions centrales des amas globulaires et des noyaux galactiques émettent des radiations

très intenses ; les astronomes sont donc de plus en plus enclins à penser qu'il y a des trous noirs au centre de ces ensembles d'étoiles.

De fait, on a détecté une source compacte et puissante de micro-ondes au centre de notre propre galaxie. Est-ce un trou noir ? Quelques astronomes pensent que oui, et que ce trou noir aurait la masse de cent millions d'étoiles, soit un millième de la masse totale de la Galaxie. Son diamètre serait cinq cents fois celui du Soleil, de l'ordre de celui d'une géante rouge. Il serait assez grand pour briser des étoiles par effet de marée, ou pour les avaler d'un seul morceau si elles s'approchaient assez vite.

Il semble maintenant que de la matière pourrait s'échapper d'un trou noir, mais pas d'une manière normale. Le physicien anglais Stephen Hawking, en 1970, a montré que le contenu énergétique d'un trou noir pourrait de temps en temps produire une paire de particules subatomiques, dont l'une pourrait s'échapper. Ainsi, le trou noir *s'évaporerait.* De gros trous noirs, de la taille d'une étoile, mettraient un temps inimaginable à s'évaporer de cette manière : des milliards de milliards de fois l'âge de l'Univers.

Mais la vitesse d'évaporation serait plus grande pour des trous noirs plus petits. Un *mini-trou noir,* pas plus massif qu'une planète ou un astéroïde (ce genre de petits objets peuvent exister si leur matière est assez dense, c'est-à-dire comprimée sous un volume assez petit) s'évaporerait assez vite pour émettre une quantité appréciable de rayons X. De plus, à mesure que l'évaporation diminuerait sa masse, la cadence d'évaporation augmenterait, et donc aussi l'émission X. Enfin, une fois devenu assez petit, le trou noir exploserait, en donnant une brève émission de rayons gamma caractéristiques.

Mais qu'est-ce qui pourrait comprimer de la matière jusqu'à la densité prodigieuse nécessaire pour qu'un mini-trou noir se forme ? Les étoiles massives peuvent être comprimées par leur propre champ gravitationnel, mais ce phénomène ne peut pas fonctionner avec un objet de la taille d'une planète, et celui-ci nécessiterait de surcroît, pour devenir un trou noir, une densité plus grande encore qu'une étoile normale.

En 1971, Hawking a suggéré que les mini-trous noirs auraient pu se former au moment du big bang, dans des conditions beaucoup plus extrêmes que tout ce qui a existé depuis. Certains de ces mini-trous auraient pu avoir une taille telle qu'il leur ait fallu tout le temps qui s'est écoulé depuis, soit quinze milliards d'années, pour s'évaporer suffisamment et arriver à l'explosion. Les astronomes pourraient ainsi déceler les émissions de rayons gamma qui serviraient à prouver leur existence.

Cette théorie est attrayante, mais jusqu'ici aucun résultat expérimental n'est venu la confirmer.

L'ESPACE « VIDE »

S'il y a dans l'Univers des objets surprenants, des surprises nous attendent aussi dans le soi-disant vide qui existe entre les étoiles. Le fait que ce « vide » ne soit pas vide du tout pose beaucoup de problèmes aux astronomes qui veulent faire des observations à distances relativement courtes.

D'une certaine manière, la galaxie que nous avons le plus de mal à observer est la nôtre. Tout d'abord, nous sommes enfermés à l'intérieur, alors que nous avons une vue d'ensemble des autres. C'est comme si on essayait de voir l'ensemble d'une ville du toit d'un petit bâtiment plutôt que d'un avion. De plus, nous sommes loin du centre, et, comble de malchance, dans un bras en spirale qui est bourré de poussières. Autrement dit, nous sommes sur le toit d'un petit immeuble de banlieue, un jour de brouillard.

Dans le meilleur des cas, l'espace qui sépare les étoiles n'est pas un vide parfait. Il y a un gaz très ténu à travers tout l'espace interstellaire à l'intérieur d'une galaxie. Les raies spectrales dues à ce *gaz interstellaire* ont été détectées en 1904 par l'astronome allemand Johannes Franz Hartmann. Dans la banlieue d'une galaxie, la concentration de gaz et de poussières devient beaucoup plus forte. Nous voyons autour des galaxies les plus proches le brouillard noir qui marque leurs confins.

En fait, nous voyons les nuages de poussières de notre propre galaxie, en négatif, sous l'aspect des régions sombres de la Voie lactée. Les exemples les plus connus sont la Nébuleuse de la Tête de Cheval, qui découpe sa silhouette noire sur un fond de millions d'étoiles et, encore plus spectaculaire, le Sac de Charbon dans la constellation de la Croix du Sud, un nuage de poussières de trente années-lumière de diamètre situé à quatre cents années-lumière environ de nous.

Si les nuages de gaz et de poussières nous empêchent de voir directement les bras en spirale de la Galaxie, ils ne les cachent pas au spectroscope. Les atomes d'hydrogène dans les nuages sont ionisés (séparés en particules subatomiques chargées électriquement) par les radiations de haute énergie venant des étoiles brillantes de population I situées dans les bras. A partir de 1951, l'astronome américain William Wilson Morgan a suivi des traînées d'hydrogène ionisé, dessinant les lignes marquées par les géantes bleues, c'est-à-dire les bras en spirale de la Galaxie. Leur spectre était analogue à celui des bras de la galaxie d'Andromède.

La plus proche de ces traînées d'hydrogène ionisé inclut les géantes bleues de la constellation d'Orion, et on appelle par conséquent cette traînée le « bras d'Orion ». Notre système solaire fait partie de ce bras. On a repéré de la même manière deux autres bras. L'un se trouve plus loin que le nôtre du centre de la Galaxie, et comprend les étoiles géantes de la constellation Persée (c'est le « bras de Persée »). L'autre est plus près du centre de la Galaxie, et contient les nuages brillants de la constellation du Sagittaire (c'est le « bras du Sagittaire »). Chacun de ces bras semble avoir à peu près dix mille années-lumière de long.

Ensuite, un instrument plus puissant encore est venu prendre la relève : la radio. Non seulement elle pouvait passer à travers les nuages sombres, mais elle pouvait leur faire raconter leur histoire – dans leur propre langage. Ce fut le résultat du travail d'un astronome hollandais, Hendrik Christoffel Van de Hulst. En 1944, la Hollande était sous la botte nazie, et les observations astronomiques étaient presque impossibles. Van de Hulst, réduit à des travaux purement théoriques, se mit à étudier les caractéristiques des atomes d'hydrogène, qui forment l'essentiel de la matière des nuages interstellaires.

Il annonça que, de temps en temps, les collisions entre ces atomes pouvaient modifier leur état énergétique, et émettre des ondes radio très

faibles. Cela pouvait arriver à un atome d'hydrogène donné une fois en onze millions d'années. Mais, étant donné leur abondance dans l'espace interstellaire, il y en aurait tout de même assez en train d'émettre à un moment donné pour que l'émission soit toujours détectable.

D'après les calculs de Van de Hulst, la longueur d'onde de cette émission devait être de vingt et un centimètres. Et de fait, les nouvelles techniques radio mises au point après la guerre devaient permettre, dès 1951, à Edward Mills Purcell et Harold Irving Ewen, de l'université de Harvard, de détecter ce « chant de l'hydrogène ».

En suivant les amas d'hydrogène émetteurs de ces ondes de vingt et un centimètres, les astronomes ont pu tracer les bras en spirale et les suivre sur de longues distances – dans la plupart des cas tout autour de la Galaxie. On trouva de nouveaux bras, et les cartes des concentrations d'hydrogène montrent maintenant plus d'une demi-douzaines de traînées.

De plus, le chant de l'hydrogène racontait quelque chose de ses mouvements. Comme toutes les ondes, cette radiation est soumise à l'effet Doppler-Fizeau. Cela permet aux astronomes de mesurer la vitesse des nuages d'hydrogène en mouvement, et par conséquent d'étudier, entre autres choses, la rotation de notre galaxie. Cette nouvelle méthode a confirmé que notre galaxie avait une période de rotation – à notre distance du centre – de deux cents millions d'années.

Dans le domaine scientifique, chaque découverte ouvre des portes menant à de nouveaux mystères. Et les progrès les plus importants viennent de l'inattendu – de la découverte qui renverse les notions admises. Par exemple, actuellement, l'étude radio des concentrations d'hydrogène au centre de la Galaxie a révélé un phénomène mystérieux. L'hydrogène semble en expansion, bien que confiné dans le plan équatorial de la Galaxie. Cette expansion en elle-même est surprenante, parce qu'aucune théorie ne permet d'en rendre compte. Et si l'hydrogène est en expansion, pourquoi ne s'est-il pas complètement dissipé durant la longue existence de la Galaxie ? Est-ce un signe que peut-être, il y a une dizaine de millions d'années, le centre de la Galaxie a explosé, comme le suppose Oort, et comme l'a fait beaucoup plus récemment celui de M 82 ? D'autre part, le disque d'hydrogène n'est pas parfaitement plan. Il se courbe vers le bas à un bout de la Galaxie et vers le haut à l'autre bout. Pourquoi ? On n'a encore proposé aucune explication satisfaisante.

L'hydrogène n'est pas – ne devrait pas être – le seul émetteur d'ondes radio. Tout atome, toute combinaison d'atomes, est capable d'émettre des radiations caractéristiques, ou d'absorber de façon caractéristique certaines des radiations émises par un « fond ». Naturellement, les astronomes ont recherché les « signatures » d'atomes moins communs que l'hydrogène.

Presque tout l'hydrogène existant est d'une variété particulièrement simple appelée *hydrogène 1.* Il en existe une forme plus complexe, appelée *deutérium* ou *hydrogène 2.* On a donc passé au peigne fin les ondes radio venues de divers points du ciel, à la recherche de la longueur d'onde prévue par la théorie. On l'a trouvée en 1966, et il semble que la quantité d'hydrogène 2 présente dans l'Univers représente à peu près 5 % de la quantité d'hydrogène 1.

Après les deux variétés d'hydrogène, les composants les plus répandus de l'Univers sont l'hélium et l'oxygène. Un atome d'oxygène peut se combiner à un atome d'hydrogène pour former un *groupe hydroxyle.* Cette

combinaison ne serait pas stable sur la Terre, parce qu'elle est extrêmement active, et se combinerait avec à peu près n'importe quel atome ou n'importe quelle molécule qu'elle rencontrerait. En particulier, le groupe hydroxyle se combinerait avec un second atome d'hydrogène pour former une molécule d'eau. Mais dans l'espace interstellaire, les atomes sont tellement séparés que les collisions sont très espacées, et un groupe hydroxyle, une fois formé, pourrait survivre longtemps, comme l'a fait remarquer en 1953 l'astronome soviétique I.S. Shklovskii.

D'après les calculs, un tel groupe hydroxyle émettrait ou absorberait quatre longueurs d'onde radio particulières. En octobre 1963, deux d'entre elles étaient détectées par une équipe d'ingénieurs radio au laboratoire Lincoln du MIT.

Étant quelque dix-sept fois plus massif qu'un atome d'hydrogène, le groupe hydroxyle est moins vif, et se déplace seulement au quart de la vitesse de l'atome d'hydrogène, à une température donnée. Comme le mouvement rend les raies plus floues, celles du groupe hydroxyle sont plus nettes que celles de l'hydrogène. Ses décalages Doppler sont plus faciles à déterminer, et si un nuage contient de l'hydroxyle, il est plus facile de dire s'il s'approche ou s'il s'éloigne.

Pour les astronomes, le plaisir de trouver une combinaison de deux atomes dans l'immensité qui sépare les étoiles n'était pas une complète surprise. Pourtant, c'est sans grand espoir qu'ils se mirent à chercher d'autres combinaisons. Les atomes sont si loin les uns des autres dans l'espace interstellaire qu'il semblait y avoir peu de chances que plus de deux atomes se rencontrent et s'assemblent d'une manière assez durable pour former une combinaison. L'idée que des atomes plus rares que l'oxygène (comme ceux du carbone et de l'azote, qui sont les suivants par ordre d'abondance) puissent intervenir dans des combinaisons semblait tout à fait invraisemblable.

Les vraies surprises ont commencé en 1968. En novembre de cette année, on a identifié la signature de la molécule d'eau. Cette molécule est formée de deux atomes d'hydrogène et d'un atome d'oxygène – trois atomes en tout. Le même mois, chose plus étonnante encore, on détectait des molécules d'ammoniac, comportant quatre atomes (trois d'hydrogène et un d'azote).

En 1969, c'était le tour d'une autre molécule de quatre atomes, celle du formaldéhyde (H_2CO), contenant un atome de carbone.

En 1970, plusieurs découvertes se sont succédé, y compris celle d'une molécule de cinq atomes, dont une chaîne de trois atomes de carbone, le cyanoacétylène (HCCCN), et une de six atomes, l'alcool méthylique (CH_3OH).

Le record devait être battu en 1971 avec les sept atomes du méthylacétylène (CH_3CCH)... et en 1982, on avait homologué une molécule de treize atomes, dont une chaîne de onze carbones ($HC_{11}N$).

Les astronomes avaient découvert une nouvelle branche de leur science : l'*astrochimie*.

On ne sait pas comment les atomes peuvent se rencontrer pour former des molécules aussi compliquées, ni comment celles-ci, une fois formées, peuvent survivre au flot de radiations de haute énergie déversé par les étoiles, et dont on s'attendrait normalement à ce qu'il les mette en pièces. Il est possible que ces molécules se forment dans des régions de l'espace moins vides que nous le croyons, peut-être dans des régions où des nuages

de poussières s'épaississent en s'engageant dans le processus de formation des étoiles.

Si c'est le cas, on détectera peut-être des molécules encore plus compliquées, et cela révolutionnerait l'idée que nous nous faisons du développement planétaire de la vie, comme nous le verrons plus loin.

I. Le système solaire

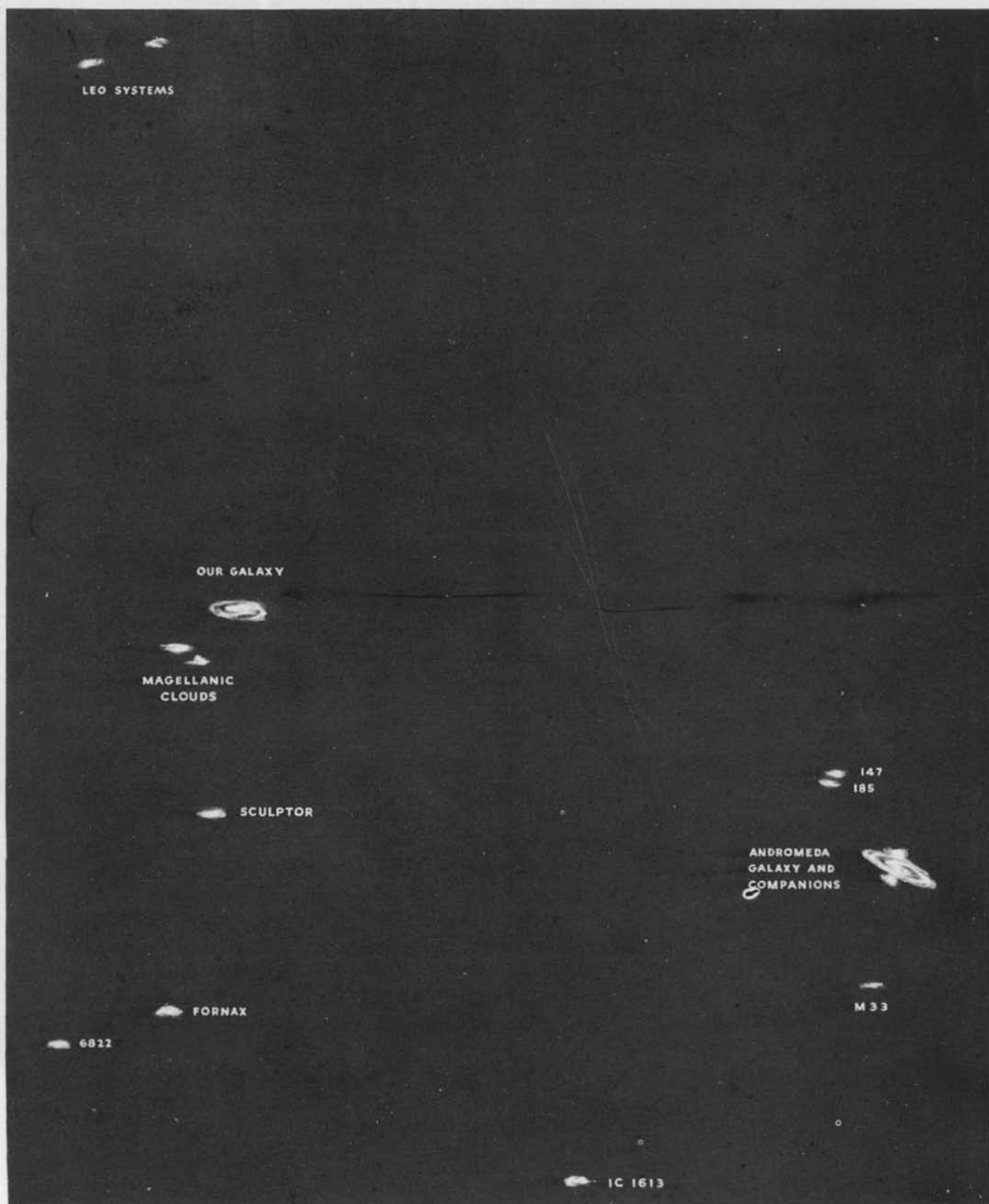


Planche I.1. Notre région de l'Univers – dessin montrant les autres galaxies de notre voisinage. (Avec l'aimable autorisation de la bibliothèque du Musée américain d'histoire naturelle.)

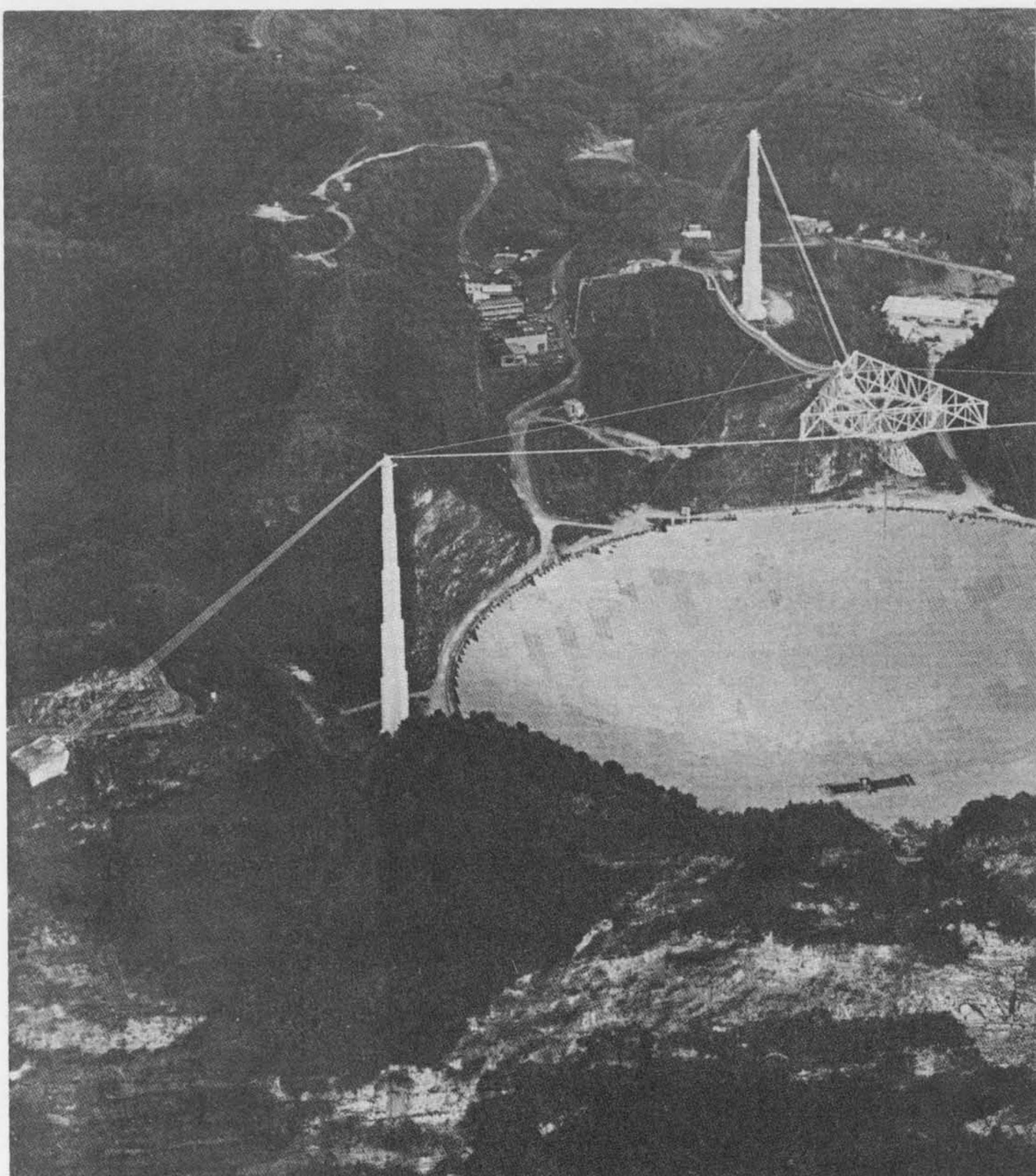


Planche I.2. Le radiotélescope de l'université Cornell. L'antenne de ce radiotélescope, situé à Arecibo, Porto Rico, a un diamètre de 300 mètres, et elle est suspendue dans une cuvette naturelle. (Avec l'aimable autorisation de l'observatoire d'Arecibo, Centre national d'astronomie et d'ionosphère, université de Cornell.)

Planche I.3. La comète de Halley, photographiée le 4 mai 1910, avec une pose de 40 minutes. (Avec l'aimable autorisation de l'observatoire de Yerkes, Wisconsin.)

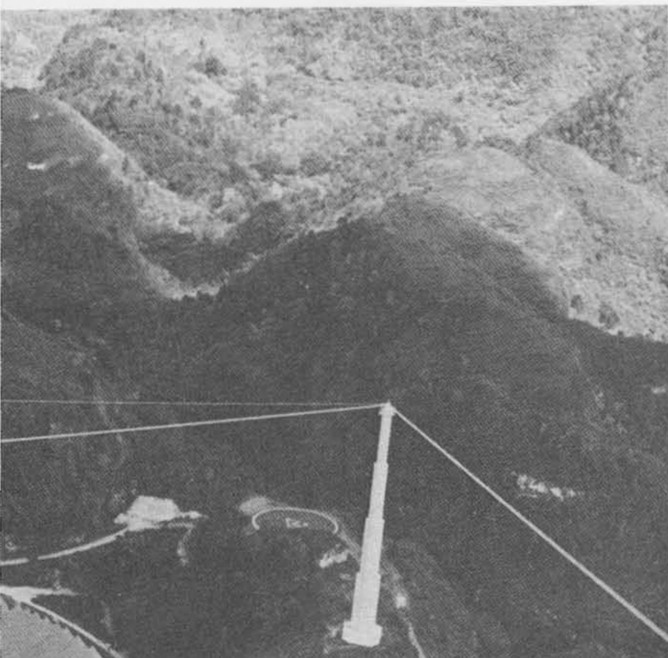




Planche I.4. Une galaxie spirale vue de face : la Nébuleuse des Chiens de Chasse. (Avec l'aimable autorisation de l'observatoire de Palomar, Californie.)



Planche I.5. Un amas globulaire dans les Chiens de Chasse. (Avec l'aimable autorisation de l'observatoire de Palomar, Californie.)



Planche I.6. La Nébuleuse du Crabe, le reste d'une supernova, photographiée en lumière rouge. (Avec l'aimable autorisation de l'observatoire de Palomar, Californie.)



Planche I.7. La Tête de Cheval, dans Orion, au sud de Zêta Orionis, photographiée en lumière rouge. (Avec l'aimable autorisation de l'observatoire de Palomar, Californie.)

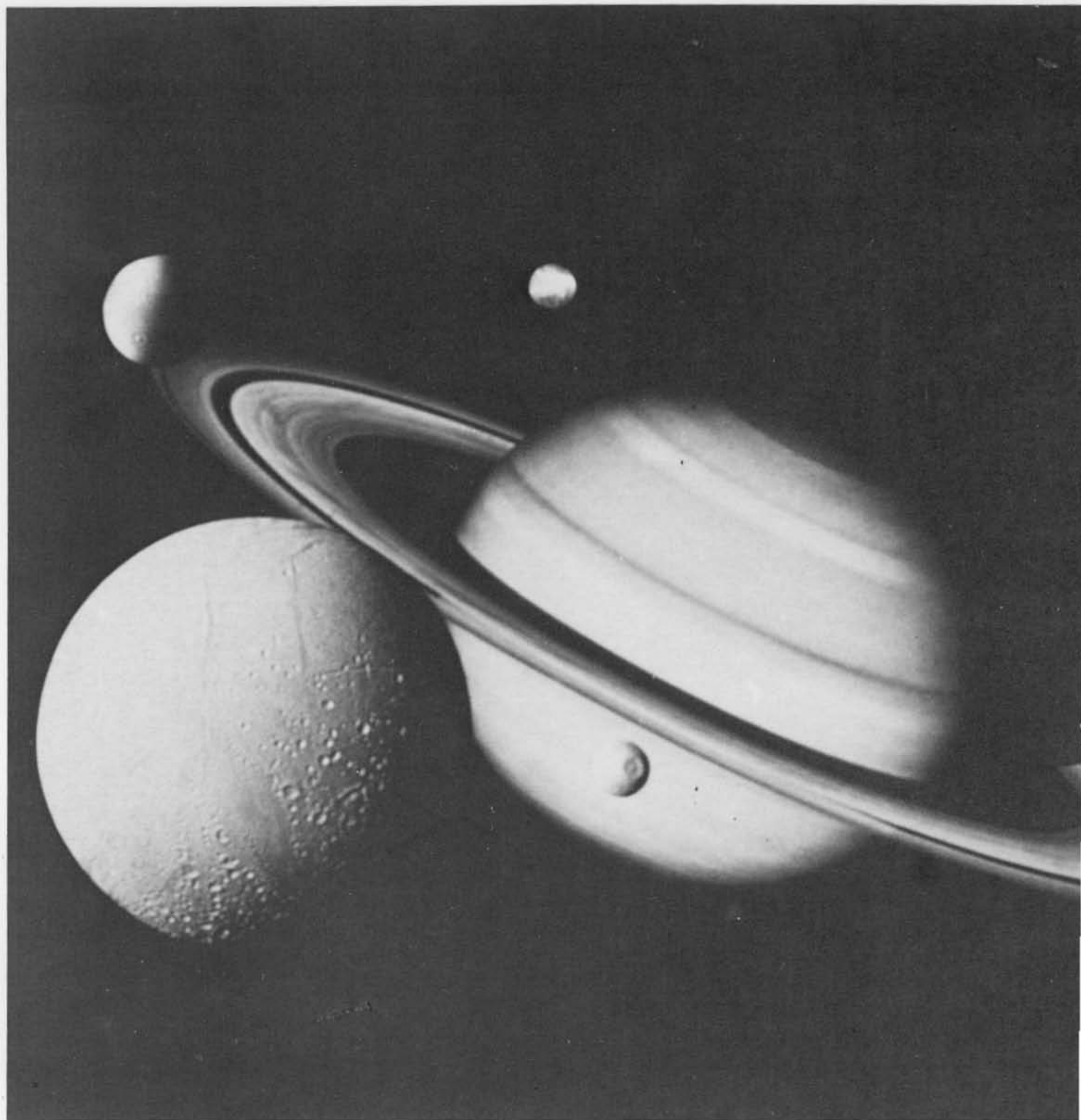
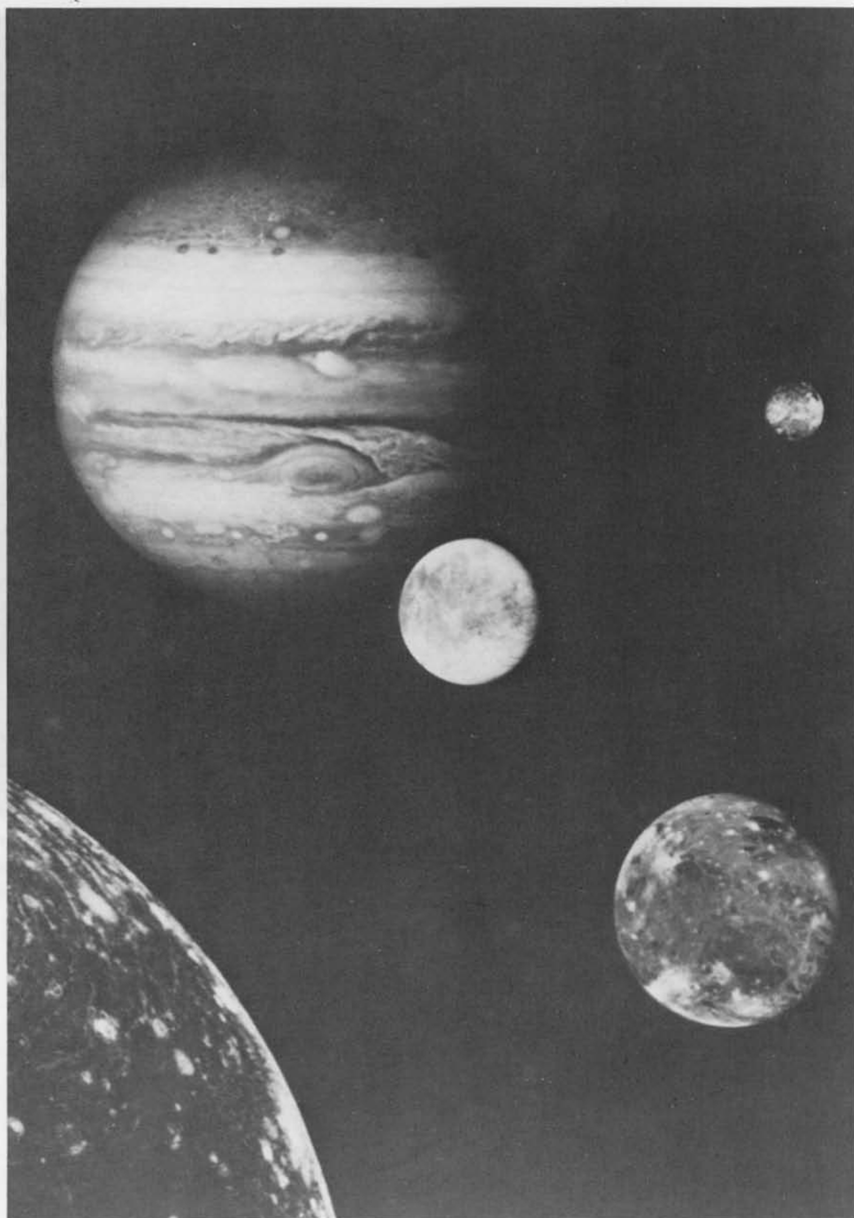


Planche I.8. Saturne et ses anneaux : montage de photographies prises par *Voyager 1* et *Voyager 2*. On voit ici tous les gros satellites de Saturne connus avant la mission Voyager (1977). Les satellites sont (*de gauche à droite et de haut en bas*) : Titan, Japet, Rhéa, Mimas et Téthys. (Avec l'aimable autorisation de la NASA.)

Planche I.9. Jupiter et ses satellites : montage de photographies prises par *Voyager 1* (1977). Les satellites galiléens sont (*de gauche à droite*) : Io, Europe, Ganymède et Callisto. (Avec l'aimable autorisation de la NASA.)



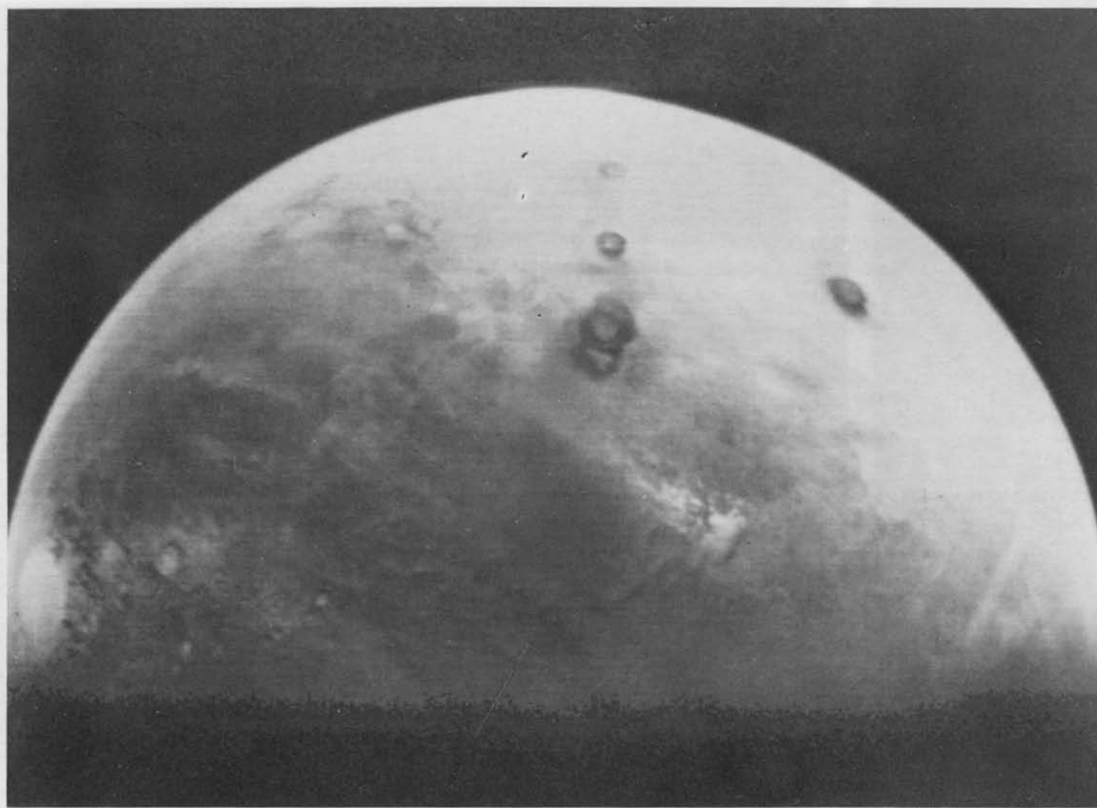


Planche I.10. Mars, photographie prise le 19 juin 1976 par *Viking 1*. On distingue bien les monts Tharsis, trois énormes volcans. Le mont Olympe, le plus grand volcan de Mars, est en haut de la photo. (Avec l'aimable autorisation de la NASA.)

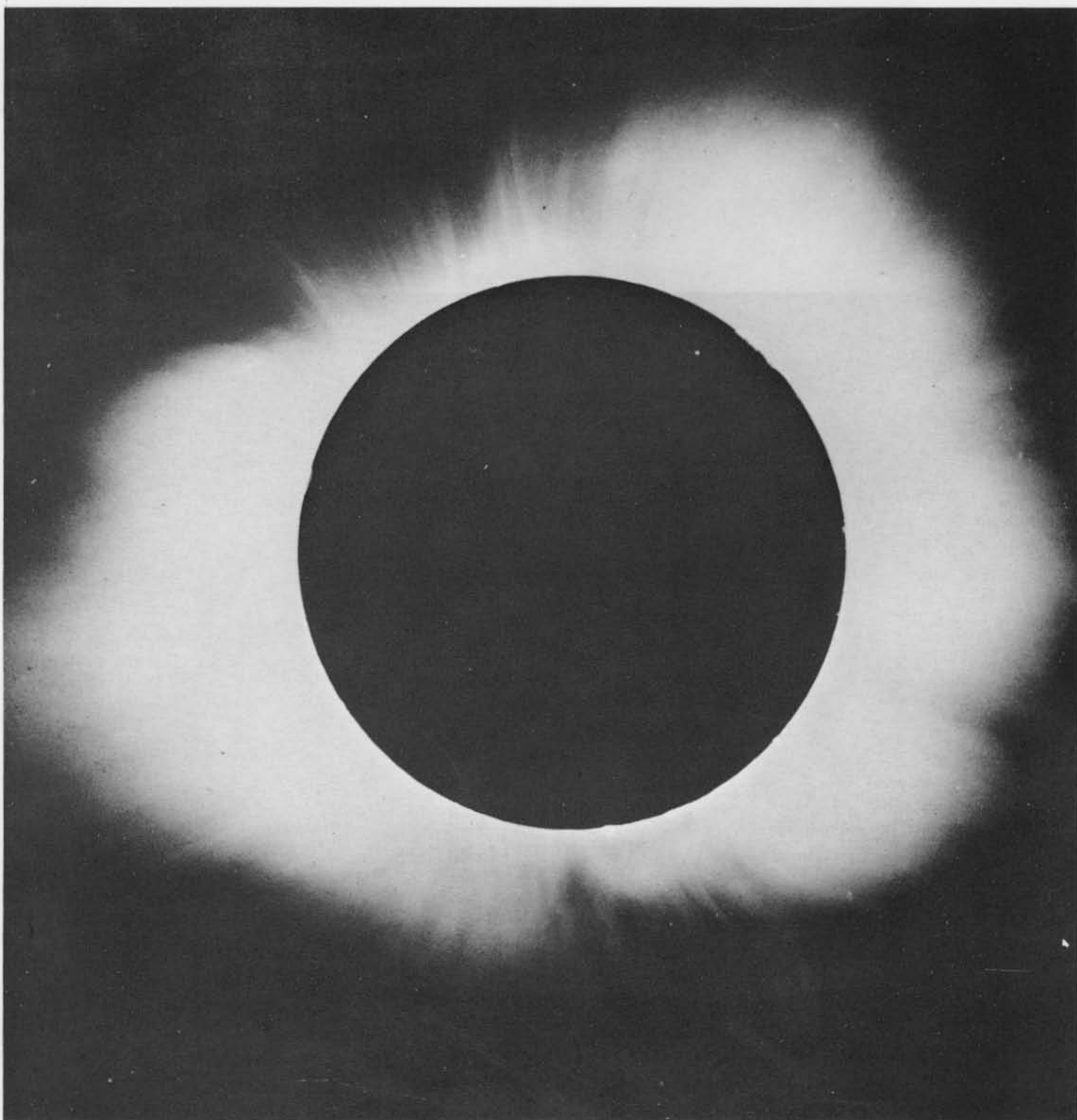


Planche I.11. La couronne solaire. (Avec l'aimable autorisation des observatoires du mont Wilson et de Las Campanas, Carnegie Institution of Washington.)

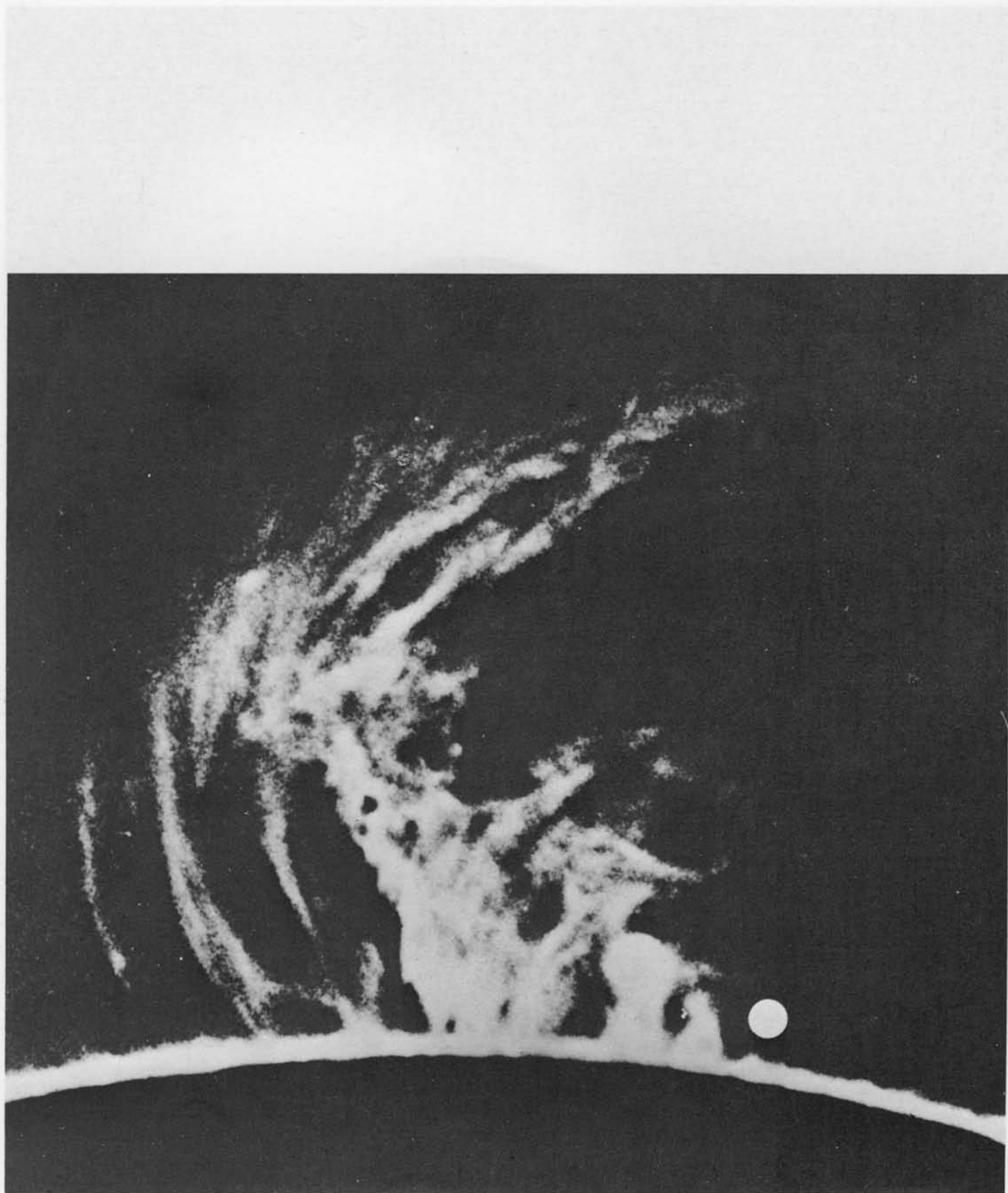


Planche I.12. Une protubérance solaire, s'étendant à 250 000 kilomètres du Soleil, photographiée à la longueur d'onde du calcium. Le cercle blanc représente la Terre à la même échelle. (Avec l'aimable autorisation des observatoires du mont Wilson et de Las Campanas, Carnegie Institution of Washington.)

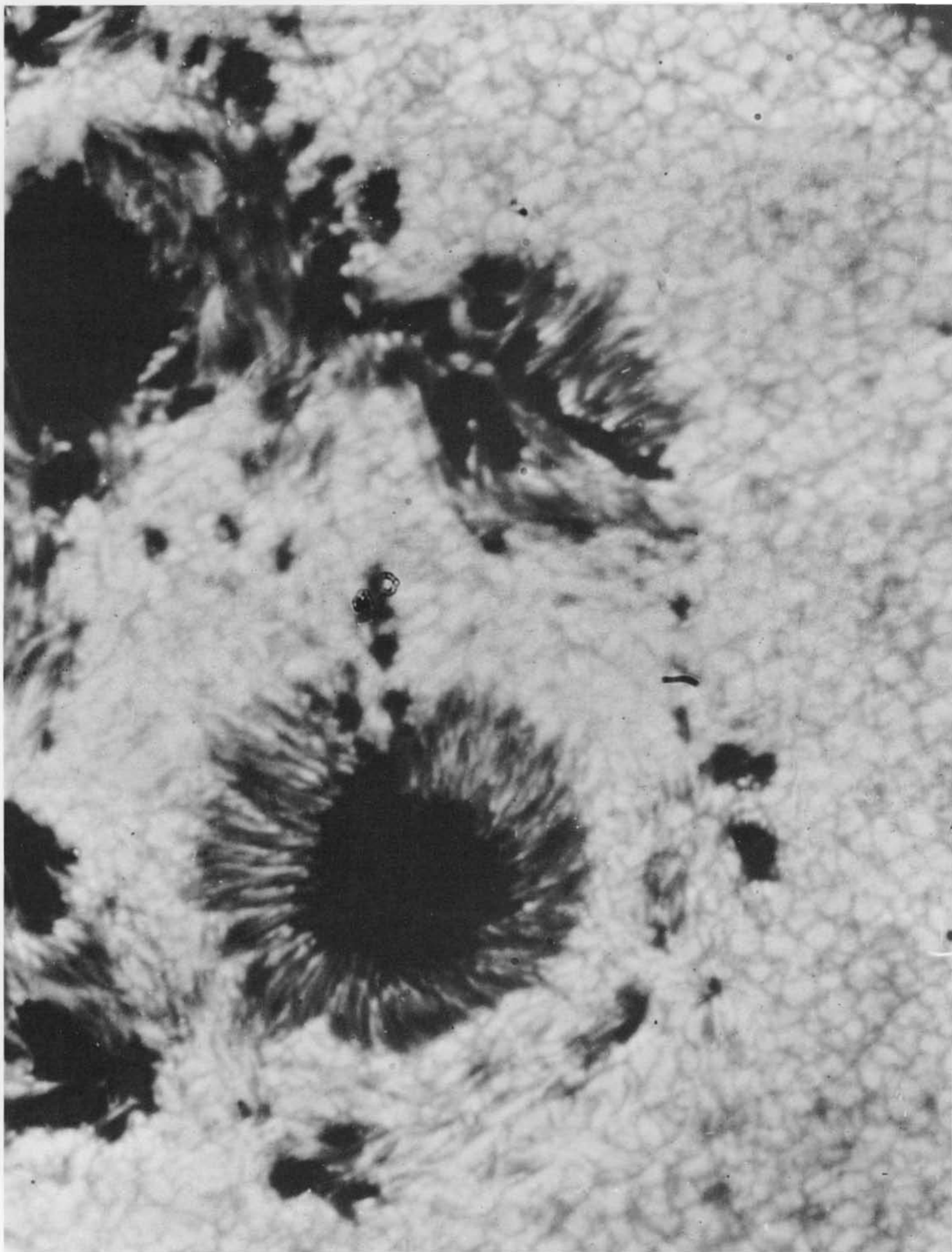


Planche I.13. Taches solaires, photographiées avec une netteté sans précédent par un ballon à 27 kilomètres d'altitude (mission américaine Stratoscope). Les taches sont formées d'un noyau sombre de gaz relativement froids, emprisonné dans un champ magnétique intense. Ce groupe de taches particulièrement actives a produit sur la Terre une aurore boréale brillante et un violent orage magnétique.



Planche I.14. Protubérances solaires. (Avec l'aimable autorisation des observatoires du mont Wilson et de Las Campanas, Carnegie Institution of Washington.)



Planche I.15. Aurore boréale. (Avec l'aimable autorisation de l'Administration nationale de l'Océan et de l'atmosphère.)

Planche I.16. Carte de la Lune. (Avec l'aimable autorisation de la NASA.)

Planche I.17. Cette vision du lever de la Terre a accueilli les astronautes d'*Apollo 8* quand ils sont ressortis de la face cachée de la Lune. Sur la Terre, à 380 000 kilomètres, le Soleil se couche en Afrique. (Avec l'aimable autorisation de la NASA.)

Chapitre 3

Le système solaire

La naissance du système solaire

Si prodigieuses et si vastes que puissent être les profondeurs inimaginables de l'Univers, nous ne pouvons pas nous perdre indéfiniment dans ses merveilles. Il nous faut retourner vers la petite famille de mondes dans laquelle nous vivons. Il nous faut retourner vers notre Soleil – une étoile parmi les centaines de milliards qui forment notre galaxie – et vers les mondes qui tournent autour de lui, parmi lesquels la Terre.

A l'époque de Newton, il était devenu possible de spéculer intelligemment sur la création de la Terre et du système solaire, en séparant ce problème de celui de la création de l'Univers dans son ensemble. Le tableau du système solaire révélait une structure dotée de certaines caractéristiques uniformes (figure 3.1) :

Toutes les planètes principales tournent autour du Soleil à peu près dans le plan de l'équateur solaire. Autrement dit, si on réalisait un modèle à trois dimensions du Soleil et de ses planètes, on s'apercevrait qu'il tient dans un moule à gâteaux très plat.

Toutes les planètes principales tournent autour du Soleil dans le même sens – le sens inverse des aiguilles d'une montre si on regarde « de dessus » (depuis un point situé dans la direction de l'Étoile polaire).

Chaque planète principale (ou presque) tourne sur son axe dans le même sens – inverse de celui des aiguilles d'une montre – qui caractérise son mouvement autour du Soleil, et le Soleil aussi tourne sur lui-même dans ce sens.

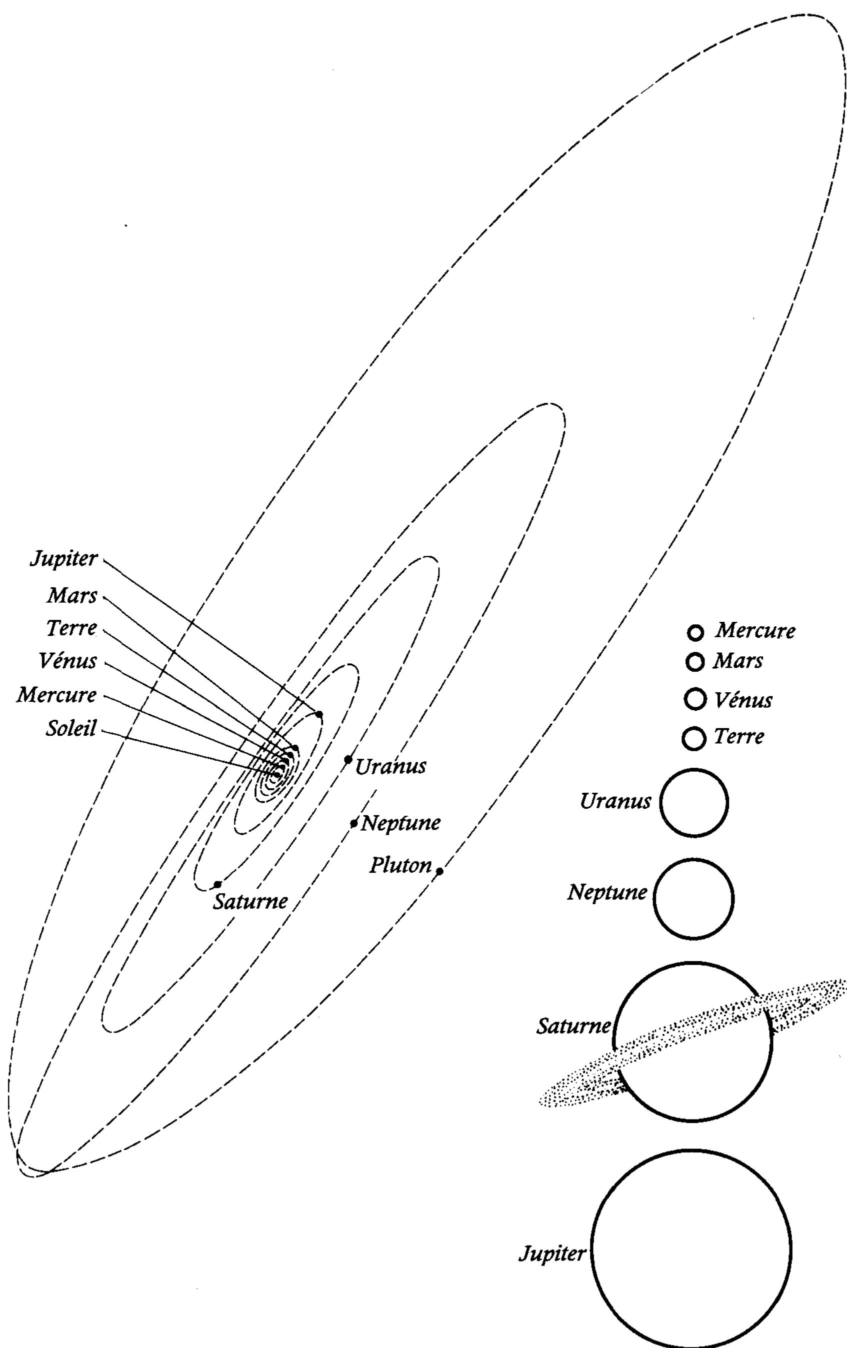


Figure 3.1. Le système solaire (représentation schématique) et la taille relative des planètes.

Les planètes ont des orbites à peu près circulaires, dont les rayons augmentent progressivement de l'une à l'autre.

Tous les satellites, à quelques exceptions près, tournent autour de leurs planètes respectives suivant des orbites presque circulaires, dans le plan de l'équateur et en sens contraire de celui des aiguilles d'une montre.

La régularité générale de ce tableau suggérait naturellement qu'un seul processus avait créé le système tout entier.

Quel peut donc être ce processus ? Toutes les théories proposées jusqu'ici se rangent en deux catégories : les « catastrophiques » et les « évolutives ». Suivant le point de vue catastrophique, le Soleil a été créé tout seul, et n'a acquis sa famille qu'à un stade relativement tardif de son histoire, à la suite de quelque événement violent. Suivant les idées évolutives, le système tout entier, Soleil et planètes, est apparu d'une façon ordonnée, dès le départ.

Au XVIII[e] siècle, quand les scientifiques étaient encore imprégnés des histoires bibliques de grands événements comme le Déluge, il était à la mode d'imaginer l'histoire de la Terre comme une série de violentes catastrophes. Alors, pourquoi pas une supercatastrophe pour mettre la chose en train ? Suivant une théorie populaire, proposée en 1745 par le naturaliste français Georges Louis Leclerc de Buffon, le système solaire se serait formé à partir des débris résultant du choc entre le Soleil et une comète.

Évidemment, Buffon voulait dire un choc entre le Soleil et un objet de masse comparable. Il appela cet objet *comète,* faute d'un autre nom. Nous savons maintenant que les comètes sont des objets très petits, entourés d'un nuage ténu de gaz et de poussières, mais l'idée de Buffon resterait valable si nous donnions un autre nom au corps heurté par le Soleil, et c'est ce qu'ont fait plus tard les astronomes.

Pour certains d'entre eux, toutefois, il semblait plus naturel, et moins fortuit, d'imaginer la naissance du système solaire comme un processus de longue durée, continu et sans catastrophes. D'une certaine façon, cela s'accorderait bien avec le tableau majestueux dressé par Newton de la loi naturelle gouvernant les mouvements de tous les mondes de l'Univers.

Newton lui-même avait suggéré que le système solaire avait pu se former à partir d'un nuage ténu de gaz et de poussières qui se serait lentement condensé sous l'effet de l'attraction gravitationnelle. Les particules se rapprochant, l'attraction serait devenue plus intense, la condensation aurait été accélérée, et finalement toute la masse se serait écrasée sur elle-même pour donner un corps compact (le Soleil), rendu incandescent par l'énergie de la contraction.

Cette idée est à la base des théories les plus répandues actuellement sur la naissance du système solaire. Mais pour répondre à des questions précises à ce sujet, il a fallu résoudre bien des problèmes épineux. Par exemple, comment un gaz très dispersé a-t-il pu se condenser sous l'effet de forces aussi faibles que le sont dans ce cas les forces de gravitation ? Une solution proposée récemment voit dans l'explosion d'une supernova le déclenchement du processus. Imaginons qu'un vaste nuage de gaz et de poussières, à peu près inchangé depuis des milliards d'années, ait dérivé par hasard dans le voisinage d'une étoile qui vient juste d'exploser en supernova. L'onde de choc de l'explosion, l'immense rafale de poussières et de gaz projetée à travers le nuage tranquille que nous avons imaginé vont le comprimer, augmentant ainsi son champ

gravitationnel, et déclenchant la condensation qui aboutira à la formation d'une étoile.

Et les planètes ? D'où viennent-elles ? Les premiers éléments de réponse sont dus à Emmanuel Kant en 1755, et indépendamment, à l'astronome et mathématicien français Pierre Simon de Laplace en 1796. C'est la représentation proposée par Laplace qui est la plus détaillée.

Selon lui, le vaste nuage de matière en train de se contracter tournait sur lui-même dès le départ. A mesure qu'il se contractait, sa vitesse de rotation augmentait, exactement comme celle d'un patineur qui ramène ses bras contre son corps en tournant sur lui-même. (Cet effet est dû à la *conservation du moment angulaire* : celui-ci étant proportionnel au produit de la vitesse par la distance au centre de rotation, quand celle-ci diminue la vitesse augmente, le produit restant constant.) Tournant de plus en plus vite, selon Laplace, le nuage commença à rejeter un anneau de matière à partir de son équateur en rotation rapide, et se débarrassa ainsi d'une partie de son moment angulaire. Le nuage restant tournait donc moins vite ; mais la contraction continuait, la vitesse augmenta à nouveau, et bientôt le nuage projeta un autre anneau de matière. Ainsi, en se condensant en Soleil, le nuage laissa derrière lui une série d'anneaux, des nuages de matière en forme de beignets. Ces anneaux, selon Laplace, se condensèrent lentement en donnant les planètes ; ce faisant, ils projetèrent eux-mêmes de petits anneaux qui formèrent leurs satellites.

Ce point de vue faisant naître le système solaire d'un nuage, ou nébuleuse, et Laplace ayant choisi comme exemple la Nébuleuse d'Andromède (on ignorait à l'époque qu'elle était une vaste galaxie, et on y voyait un nuage de poussières et de gaz tournant sur lui-même), on désigne généralement cette théorie sous le nom d'*hypothèse de la nébuleuse de Laplace.*

Cette hypothèse semblait rendre fort bien compte des caractéristiques principales du système solaire – et même de certains détails. Par exemple, les anneaux de Saturne pouvaient être des « anneaux-satellites » qui ne s'étaient pas condensés (en les rassemblant, on obtiendrait un satellite de taille respectable). De même, les astéroïdes qui tournent autour du Soleil entre l'orbite de Mars et celle de Jupiter pouvaient provenir des morceaux d'un anneau qui ne se seraient pas rassemblés en une seule planète. Les théories d'Helmholtz et de Kelvin, attribuant l'énergie libérée par le Soleil à sa lente contraction, semblaient aussi s'inscrire fort bien dans le tableau dessiné par Laplace.

L'hypothèse de la nébuleuse a régné en maîtresse pendant la plus grande partie du XIXe siècle. Mais bien avant que celui-ci s'achève, de sérieux doutes ont commencé à apparaître. En 1859, James Clerk Maxwell analysa mathématiquement les anneaux de Saturne, et montra qu'un anneau de matière gazeuse rejeté par un corps quelconque ne peut donner par condensation qu'une multitude de petits objets comme ceux qui forment les anneaux de Saturne ; il ne peut absolument pas former un corps solide, parce que les forces de gravitation fragmentent l'anneau avant qu'il puisse se condenser.

Un problème se posait aussi à propos du moment angulaire. Les planètes, qui forment à elles seules à peine un millième de la masse du système solaire, possèdent 98 % de son moment angulaire total ! Jupiter à lui seul porte 60 % du moment angulaire total du système solaire. Le Soleil, par conséquent, ne garde qu'une fraction très petite du moment angulaire que

possédait le nuage initial. Comment la quasi-totalité de ce moment angulaire a-t-elle pu se trouver transférée dans les petits anneaux laissés derrière elle par la nébuleuse ? Le mystère est d'autant plus grand que les systèmes de satellites de Jupiter et de Saturne, qui ressemblent à des miniatures du système solaire, ont probablement été formés de la même façon que lui ; or, dans les deux cas, le corps planétaire central garde la plus grande partie du moment angulaire.

Vers 1900, l'hypothèse de la nébuleuse était si complètement abandonnée que toute idée d'un processus évolutif semblait discréditée. Le rideau allait se lever sur la résurrection des théories catastrophiques. En 1905, les Américains Thomas Chrowder Chamberlin et Forest Ray Moulton, utilisant un nom plus satisfaisant que la *comète* de Buffon, annoncèrent que les planètes étaient nées quand une étoile avait frôlé le Soleil. Cet accident avait arraché des deux soleils des nuages de gaz, et ceux qui restaient au voisinage de notre Soleil s'étaient condensés ensuite en *planéticules,* qui se rassemblèrent ensuite en planètes. C'est l'*hypothèse planéticulaire.* Quant au problème du moment angulaire, les Anglais James Hopwood Jeans et Harold Jeffreys proposèrent en 1918 une *hypothèse de la marée,* suivant laquelle l'attraction gravitationnelle sur l'étoile qui avait frôlé le Soleil avait donné un « effet » aux masses de gaz (comme on « coupe » une balle), ce qui leur avait donné un moment angulaire important. Si cette théorie catastrophique était juste, il y aurait extrêmement peu de systèmes planétaires. Les étoiles sont si espacées que les collisions entre elles sont dix mille fois moins fréquentes que les supernovae, qui ne sont elles-mêmes pas fréquentes... On estime que, depuis la formation de la Galaxie, il se serait produit une dizaine de rencontres du type qui, suivant cette théorie, peut donner naissance à un système solaire.

De toute façon, ces premiers scénarios catastrophiques ne devaient pas résister à l'épreuve de l'analyse mathématique. Comme le démontra Russel, de telles collisions laisseraient les planètes des milliers de fois plus loin du Soleil qu'elles ne le sont en réalité. Des tentatives ultérieures pour reconstituer la théorie en imaginant toutes sortes de vraies collisions, au lieu du « frôlement », ne devaient pas connaître un grand succès. Dans les années trente, Lyttleton a évoqué la possibilité d'une collision de trois étoiles, et plus tard Hoyle a imaginé un compagnon du Soleil explosant en supernova et laissant derrière lui un héritage de planètes. Mais en 1939, l'astronome américain Lyman Spitzer a montré que la matière éjectée par le Soleil, dans n'importe quelle circonstance, serait si chaude qu'au lieu de se condenser en planéticules elle se disperserait en gaz ténu. Ceci semblait sonner le glas de toutes les théories catastrophiques (bien qu'en 1965 un astronome anglais, M.M. Woolfson, ait suggéré que le Soleil avait pu tirer ses matériaux planétaires d'une étoile froide très peu dense, ce qui permettait d'éviter ces très hautes températures).

Aussi, la théorie planéticulaire étant arrivée dans cette impasse, les astronomes revinrent aux théories évolutives, et examinèrent à nouveau l'hypothèse de la nébuleuse de Laplace.

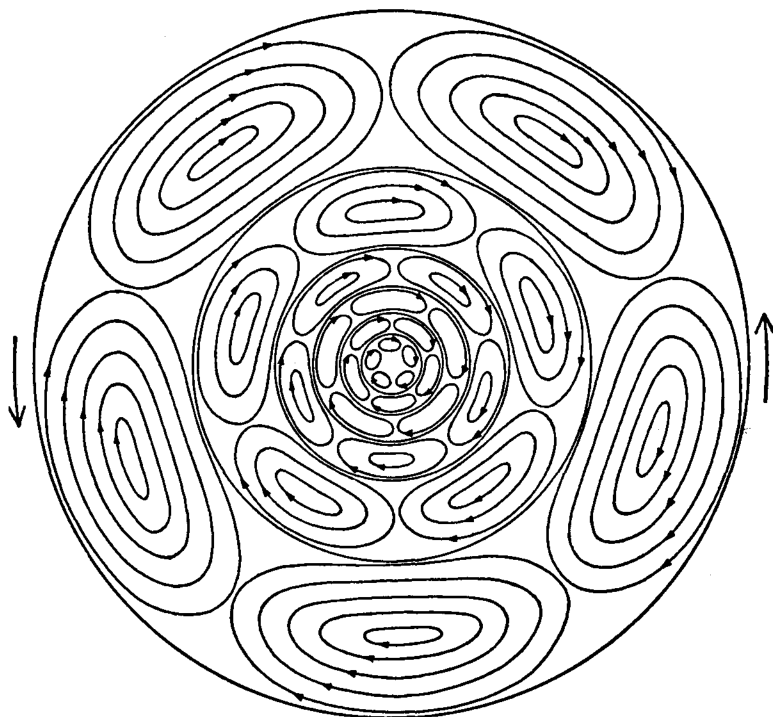
A ce moment-là, leurs connaissances sur l'Univers s'étaient considérablement étendues. Ils devaient désormais rendre compte de la formation des galaxies, qui mettent en jeu des nuages de gaz et de poussières énormément plus importants que celui où Laplace voyait le père du système solaire. On réalisa que ces gigantesques amas de matière, sous l'effet de la

turbulence, se séparaient en un grand nombre de tourbillons, dont chacun pouvait donner naissance, par condensation, à un système particulier.

En 1944, l'astronome allemand Carl F. von Weizsäcker analysa à fond cette idée. Suivant ses calculs, les plus vastes de ces tourbillons contenaient assez de matière pour former des galaxies. Au cours de la contraction turbulente de ces tourbillons, il apparaissait des tourbillons d'une échelle plus réduite. Chacun de ces tourbillons avait la taille nécessaire pour donner naissance à un système solaire (avec un ou plusieurs soleils). A la limite du tourbillon solaire, des tourbillons d'une taille encore inférieure pouvaient donner naissance à des planètes. Ainsi, aux limites où plusieurs tourbillons se rencontraient comme les roues dentées d'un engrenage, des poussières s'aggloméraient, formant des planéticules, puis des planètes (figure 3.2).

La théorie de Weizsäcker, par elle-même, ne résolvait pas mieux le problème du moment angulaire des planètes que la version beaucoup plus simple proposée par Laplace. L'astrophysicien suédois Hannes Alfven fit intervenir le champ magnétique du Soleil. Lors de la rotation rapide du jeune Soleil, son champ magnétique le freina et transféra aux planètes

Figure 3.2. L'origine du système solaire, selon le modèle de Carl F. von Weizsäcker. Suivant sa théorie, le grand nuage initial se sépara en tourbillons et en sous-tourbillons, qui se condensèrent ensuite pour donner le Soleil, les planètes et leurs satellites.



le moment angulaire. Développée ensuite par Hoyle, cette idée a fait de la théorie de Weizsäcker, modifiée pour inclure les forces magnétiques à côté de la gravitation, l'explication la plus satisfaisante à ce jour de l'origine du système solaire.

Le Soleil

De toute évidence, le Soleil est la source de la lumière, de la chaleur et de la vie sur la Terre, et l'homme préhistorique en avait sans doute fait un dieu. Le pharaon Akhenaton, monté sur le trône d'Égypte en 1379 av. J.-C., fut le premier monothéiste que nous connaissions, et pour lui le dieu unique était le Soleil. Au Moyen Age, le Soleil était le symbole de la perfection, et bien qu'on ne le considérât pas lui-même comme un dieu, il représentait certainement la perfection du Tout-Puissant.

Les Grecs de l'Antiquité furent les premiers à avoir eu une idée de la distance réelle qui nous en sépare : les observations d'Aristarque montraient que cette distance était au moins de plusieurs millions de kilomètres, et sa taille apparente laissait supposer que le Soleil était plus gros que la Terre. A elle seule, toutefois, cette taille n'était pas vraiment impressionnante, car on considérait le Soleil comme une simple boule de lumière, sans substance.

C'est seulement à l'époque de Newton qu'il devint évident que le Soleil devait être non seulement plus vaste, mais aussi beaucoup plus massif que la Terre, et que celle-ci tournait autour de lui parce qu'elle était liée par le champ gravitationnel intense du Soleil. Nous savons maintenant que le Soleil est à 149 millions de kilomètres de la Terre, et que son diamètre vaut 1 400 000 kilomètres, soit environ cent dix fois celui de la Terre. Sa masse vaut 330 000 fois celle de la Terre, et 745 fois celle de toute la matière planétaire du système. Autrement dit, le Soleil contient à lui tout seul 99,86 % de toute la matière présente dans le système solaire, ce qui en fait sans conteste le membre le plus important...

Mais ne nous laissons pas impressionner par une simple question de taille : ce n'est certainement pas un corps parfait, si nous entendons par là – comme les érudits du Moyen Age – un corps uniformément brillant et sans tache.

A la fin de l'année 1610, Galilée se servit de sa lunette pour observer le Soleil, quand son éclat est atténué juste avant le coucher, et distingua chaque jour des taches sombres sur le disque solaire. Il remarqua que ces taches se déplaçaient d'une façon continue à travers le disque, et que leur forme s'aplatissait quand elles s'approchaient du bord ; il en conclut qu'elles faisaient partie de la surface du Soleil, et que celui-ci tournait autour de son axe en un peu plus de vingt-cinq jours terrestres.

Bien entendu, les découvertes de Galilée soulevèrent une opposition considérable ; en effet, pour les tenants des idées antérieures, elles étaient blasphématoires. Un astronome allemand, Christoph Scheiner, qui observa aussi les taches, suggéra qu'elles ne faisaient pas partie du Soleil, mais étaient de petits corps qui tournaient autour de lui et se détachaient en noir sur le fond lumineux du disque. Dans ce débat, toutefois, Galilée finit par l'emporter.

En 1774 un astronome écossais, Alexander Wilson, remarqua qu'une grande tache voisine du bord du Soleil, vue par conséquent de côté, semblait concave, comme s'il s'agissait d'un cratère à la surface du Soleil. En 1795 Herschel revint sur ce point, et suggéra que le Soleil était un corps sombre et froid, entouré d'un manteau de gaz embrasés. Selon ce point de vue, les taches étaient des trous à travers lesquels on apercevait le corps sombre qui se trouvait sous le manteau. Herschel se laissa aller à imaginer que ce corps froid pouvait même être habité par des êtres vivants. (Un savant de premier plan propose parfois des idées audacieuses, raisonnables certes à la lumière des connaissances de son temps, mais qui se révèlent fausses quand les connaissances sur le sujet se développent.)

En réalité, les taches solaires ne sont pas vraiment noires. Ce sont des régions de la surface qui sont moins chaudes que le reste, et paraissent sombres par comparaison. Mais si Mercure ou Vénus viennent s'interposer entre le Soleil et nous, elles apparaissent sur le disque comme de petits ronds *vraiment* noirs, et si ce petit rond s'approche d'une tache, on se rend bien compte que la tache, elle, n'est pas vraiment noire.

Toutefois, des idées totalement fausses peuvent avoir leur utilité, et celles d'Herschel ont attisé l'intérêt éveillé par les taches solaires.

Mais la découverte importante est due à un pharmacien allemand, Heinrich Samuel Schwabe, amateur passionné d'astronomie. Travaillant toute la journée, il ne pouvait pas passer ses nuits en observation. Il chercha donc un sujet d'étude diurne, et décida d'observer le disque solaire, dans l'espoir que des planètes nouvelles, proches du Soleil, révèlent leur existence en passant devant lui.

Il commença ses observations en 1825, et ne put s'empêcher d'être frappé par les taches solaires. Au bout de quelque temps, il oublia ses planètes, et se mit à dessiner les taches solaires, qui changeaient de place et de forme d'un jour à l'autre. Il poursuivit ces observations pendant dix-sept ans, ne les délaissant que les jours où le ciel était complètement couvert.

En 1843, il put annoncer que les taches n'apparaissaient pas d'une façon désordonnée : leur apparition était cyclique. D'une année à l'autre, leur nombre augmentait, jusqu'à atteindre un maximum. Ensuite, elles devenaient de plus en plus rares, jusqu'à disparaître presque complètement ; et un nouveau cycle commençait. Nous savons maintenant que ce cycle, assez irrégulier, dure en moyenne onze ans. La découverte de Schwabe passa inaperçue (après tout, ce n'était qu'un pharmacien...) jusqu'à ce que le savant bien connu Alexandre von Humboldt mentionne ce cycle dans son livre *Kosmos* (1851), une sorte de revue générale des connaissances scientifiques.

A la même époque, l'astronome germano-écossais Johann von Lamont mesura régulièrement l'intensité du champ magnétique terrestre, et montra que celui-ci croissait et décroissait périodiquement. En 1852, un physicien anglais, Edward Sabine, remarqua que ce cycle coïncidait avec celui des taches solaires.

Il devenait donc évident que les taches solaires avaient un effet sur la Terre, et on commença à les étudier avec un intérêt très vif. On attribua à chaque année un *nombre de taches de Zurich,* suivant une formule mise au point en 1849 par un astronome suisse, Rudolf Wolf, de Zurich. (Il fut le premier à avoir remarqué que la fréquence des aurores boréales, elle aussi, suit le cycle des taches solaires.)

Les taches semblent être liées au champ magnétique solaire, et apparaître aux points d'émergence des lignes de force magnétiques. En 1908, trois siècles après la découverte des taches, G.E. Hale montra qu'un champ magnétique intense leur est associé. Quant à savoir pourquoi le champ magnétique du Soleil se comporte de cette manière, sort de la surface à tel ou tel moment et à tel ou tel endroit, augmente et diminue d'une façon à peu près cyclique, c'est l'un des mystères qui subsistent à propos du Soleil.

En 1893, l'astronome anglais Edward Walter Maunder passa en revue les premiers comptes rendus relatifs aux taches, pour mettre en ordre les résultats disponibles pendant le siècle qui suivit la découverte de Galilée. A sa grande surprise, il s'aperçut qu'il n'y avait pratiquement aucune tache signalée entre 1645 et 1715. Des astronomes de premier plan, comme Cassini, cherchèrent des taches et signalèrent qu'ils n'en trouvaient aucune. Maunder publia ses résultats en 1894, puis une seconde fois en 1922, mais son travail passa complètement inaperçu. Le cycle des taches était si bien établi qu'une période de soixante-dix ans sans taches paraissait tout à fait invraisemblable.

En 1970, l'astronome américain John A. Eddy tomba par hasard sur ces publications, les vérifia, et constata qu'il semblait bien s'être produit ce qu'on appelle dorénavant un *minimum de Maunder*. Il ne se contenta pas de refaire les recherches de Maunder, mais il étudia les comptes rendus d'observations à l'œil nu de taches particulièrement importantes, provenant des régions les plus variées du globe, y compris l'Extrême-Orient – résultats qui n'avaient pas été à la disposition de Maunder. Ces comptes rendus remontent au ve siècle av. J.-C., et donnent généralement une dizaine d'observations par siècle. Mais il y a des « trous », et l'un de ces trous correspond au minimum de Maunder.

Eddy releva également les comptes rendus relatifs aux aurores boréales, dont la fréquence suit le même cycle que les taches solaires. Il en trouva beaucoup après 1715, un certain nombre avant 1645, mais presque aucune entre ces deux dates.

De plus, quand l'activité magnétique du Soleil est grande et les taches nombreuses, la couronne solaire est très belle et pleine de « rayons » de lumière. En l'absence de taches, elle se réduit à une brume lumineuse indistincte. On peut l'observer lors des éclipses totales du Soleil. Or, si les astronomes étaient peu nombreux à se déplacer, au XVIIe siècle, pour observer les éclipses totales, tous les rapports de cette époque mentionnent une couronne dont l'aspect correspond à l'absence de taches.

Enfin, au moment où le nombre des taches passe par un maximum, la production de carbone 14 (une variété de carbone dont nous parlerons dans le chapitre suivant) est moins importante que le reste du temps. On peut mesurer la proportion de carbone 14 dans les anneaux successifs des arbres et évaluer à partir de là l'activité solaire de l'année correspondante. Ce type d'analyse confirme l'existence du minimum de Maunder et de nombreux minima de ce type dans les siècles antérieurs.

Selon Eddy, il semble y avoir eu, en cinq mille ans, une douzaine de périodes correspondant à des minima de Maunder, dont la durée était à chaque fois de cinquante à deux cents ans. Par exemple, il y en avait eu une entre 1400 et 1510.

Puisque le cycle des taches solaires a un effet sur la Terre, on peut se demander quel est l'effet des minima de Maunder. Ils pourraient correspondre à des périodes de froid. Les hivers étaient si froids en Europe

dans les dix premières années du XVIIIe siècle qu'on a appelé cette période la *petite glaciation*. De la même façon, le froid était si intense pendant le minimum de 1400-1510 que la colonie norvégienne du Groenland cessa d'exister, simplement parce que le temps ne permettait pas d'y survivre.

La Lune

Quand Copernic, en 1543, plaça le Soleil au centre du système, il ne restait plus que la Lune à reconnaître la suzeraineté de la Terre, considérée jusque-là comme le centre du monde.

La Lune fait le tour de la Terre (par rapport aux étoiles) en 27,32 jours. Elle met exactement le même temps à tourner sur elle-même. Grâce à cette égalité entre ses périodes de rotation et de révolution, la Lune tourne toujours la même face vers la Terre. Cette égalité n'est pas une coïncidence : elle résulte de la « marée » terrestre sur la Lune, comme nous le verrons bientôt.

Le temps que met la Lune à faire un tour par rapport aux étoiles est le *mois sidéral*. Mais pendant que la Lune tourne autour de la Terre, la Terre tourne autour du Soleil. Quand la Lune a fait un tour entier autour de la Terre, le Soleil s'est déplacé dans le ciel, à cause du mouvement de la Terre (qui a entraîné la Lune avec elle). La Lune doit continuer à tourner pendant deux jours et demi pour « rattraper le Soleil » et se retrouver, par rapport à lui, dans la même position qu'au début. Le temps que met la Lune à faire ce tour autour de la Terre, par rapport au Soleil, est le *mois synodique*, soit 29,53 jours.

Le mois synodique est plus important pour les hommes que le mois sidéral, car dans le mouvement de la Lune autour de la Terre, l'éclairement de la face visible change progressivement, d'une manière qui dépend de la position de la Lune par rapport au Soleil. La Lune passe ainsi à travers une succession de *phases*. Au début d'un mois synodique, la Lune se trouve juste à l'est du Soleil, et apparaît sous la forme d'un très fin croissant, visible immédiatement après le coucher du Soleil.

D'une nuit à l'autre, la Lune s'éloigne du Soleil, et le croissant s'épaissit. Il devient bientôt un demi-cercle, puis grossit encore. Quand la Lune s'est suffisamment déplacée pour se trouver dans la région du ciel opposée à celle où se trouve le Soleil, celui-ci l'éclaire « par-dessus la Terre », et toute la face visible de la Lune est éclairée : ce disque de lumière s'appelle la *pleine lune*.

Puis l'ombre commence à rogner le disque du côté où le croissant avait commencé à apparaître. Nuit après nuit, la portion éclairée de la Lune se rétrécit, jusqu'à redevenir une demi-lune, la moitié éclairée étant celle qui se trouvait dans l'ombre pour la précédente demi-lune. Finalement, la Lune se retrouve immédiatement à l'ouest du Soleil, et apparaît dans le ciel juste avant l'aube, sous la forme d'un croissant tourné en sens opposé à celui du croissant initial. Puis la Lune dépasse le Soleil, réapparaît en croissant juste après le coucher du Soleil, et toute la série de changements recommence.

Le cycle entier dure vingt-neuf jours et demi – un mois synodique : c'est la base des premiers calendriers de l'humanité.

Les hommes ont commencé par croire que la Lune grossissait et diminuait vraiment à mesure que les phases se succédaient. Ils imaginaient même, à chaque fois que le croissant réapparaissait à l'ouest après le coucher du Soleil, qu'il s'agissait littéralement d'une *nouvelle Lune,* nom que garde encore aujourd'hui cette phase de la Lune.

Mais les astronomes de l'Antiquité grecque réalisèrent bientôt que la Lune devait être une boule, que ses phases étaient dues au fait qu'elle brillait seulement là où le Soleil l'éclairait, et que ses changements de position dans le ciel par rapport au Soleil expliquaient parfaitement les changements de forme de sa partie éclairée. Ceci avait une importance considérable à l'époque. Les philosophes grecs, Aristote en particulier, tentèrent de faire une distinction entre la Terre et les corps célestes en montrant que les propriétés de la Terre étaient radicalement différentes de celles que possédaient en commun tous les objets célestes. Ainsi, la Terre était sombre, et n'émettait pas de lumière, tandis que les corps célestes en émettaient tous. Pour Aristote, ceux-ci étaient formés d'une substance qu'il appela *éther* (d'un mot grec signifiant « brillant » ou « éclatant »), substance fondamentalement différente des matériaux qui constituaient la Terre. Or, la succession des phases de la Lune montrait que celle-ci, pas plus que la Terre, n'émettait de lumière, et qu'elle ne brillait qu'en renvoyant les rayons du Soleil. Ainsi, la Lune au moins ressemblait sur ce point à la Terre.

Qui plus est, il arrivait que le Soleil et la Lune fussent si exactement opposés par rapport à la Terre que celle-ci empêchait les rayons du Soleil d'atteindre la Lune. La Lune (toujours pleine à ce moment-là) passait dans l'ombre de la Terre : c'était une éclipse de Lune.

Pour les hommes primitifs, la Lune était alors avalée par une puissance malfaisante, et devait disparaître à jamais. C'était évidemment effrayant, et l'une des premières victoires de la science fut d'être capable de prédire une éclipse et de montrer qu'il s'agissait d'un phénomène naturel, facile à comprendre. (Certains voient en Stonehenge, entre autres choses, un observatoire de l'âge de pierre, qui pouvait servir à prédire les éclipses de Lune, grâce aux déplacements du Soleil et de la Lune par rapport aux pierres, régulièrement espacées, qui forment cette structure.)

En fait, quand la Lune est un croissant fin, on peut voir le « reste de la Lune » faiblement éclairé par la « lumière cendrée ». C'est à Galilée que nous devons l'idée que la Terre, comme la Lune, doit renvoyer les rayons du Soleil et « briller », et que c'est ce « clair de Terre » qui illumine la partie obscure de la Lune. On s'en aperçoit seulement quand la partie éclairée par le Soleil est suffisamment réduite pour que sa lumière vive ne masque pas le « clair de Terre » beaucoup plus discret. Ainsi, non seulement la Lune était aussi incapable que la Terre d'émettre de la lumière, mais voilà que la Terre était capable de renvoyer les rayons du Soleil et avait aussi des phases (si on la voyait depuis la Lune).

Une autre « différence fondamentale » entre la Terre et les corps célestes était que la Terre était imparfaite et pleine de défauts, tandis que les corps célestes étaient immuables et parfaits.

A l'œil nu, seuls le Soleil et la Lune apparaissent comme quelque chose de plus que des points de lumière. Des deux, le Soleil semble bien être un cercle parfait de lumière parfaite. La Lune, par contre – même sans parler des phases – n'est pas parfaite. Quand elle est pleine et apparaît comme un cercle parfait de lumière, on voit pourtant qu'elle n'est pas

parfaite du tout. Des traces sombres sur sa surface brillante lui enlèvent toute prétention à la perfection. Les premiers hommes voyaient des dessins dans ces traces sombres, ces dessins différant d'une culture à l'autre. Le plus souvent, la fatuité de l'espèce humaine y voyait l'image d'un être humain, et on parle encore à notre époque de « l'homme de la Lune ».

C'est Galilée, en 1609, qui regarda le premier le ciel à travers une lunette. Quand il la pointa vers la Lune, il y découvrit des montagnes, des cratères et des étendues plates (qu'il prit pour des mers). Cette fois, c'était bien sûr : la Lune n'était pas un corps céleste « parfait », fondamentalement différent de la Terre, mais au contraire un monde du même type qu'elle.

Cette constatation ne suffit pas à elle seule, cependant, à démolir complètement les idées antérieures. Les Grecs avaient remarqué que plusieurs objets dans le ciel se déplaçaient par rapport aux étoiles, et que de tous ces objets, la Lune était le plus proche de la Terre (et changeait de position le plus vite). On pouvait donc supposer que la Lune, à cause de cette proximité, était en quelque sorte contaminée par les imperfections de la Terre, qu'elle avait souffert de ce voisinage. C'est seulement quand Galilée découvrit les taches du Soleil que la perfection céleste commença à être mise sérieusement en doute.

MESURER LA LUNE

Si la Lune est bien l'objet le plus proche de la Terre, à quelle distance se trouve-t-elle ? Après plusieurs autres astronomes de l'Antiquité grecque, Hipparque s'attaqua au problème, et en donna une solution très proche des résultats modernes : nous savons maintenant que la distance moyenne entre la Terre et la Lune est de 380 000 kilomètres, neuf fois et demie la circonférence terrestre.

Si l'orbite de la Lune était circulaire, cette distance serait constante. Mais l'orbite de la Lune est légèrement elliptique, et la Terre n'est pas au centre de l'ellipse, mais à l'un des foyers, qui sont excentrés. La Lune s'approche un peu de la Terre sur la moitié de son orbite, et s'en éloigne sur l'autre moitié. Au point où la distance est minimale (*périgée*), elle vaut 356 000 kilomètres, et au point où elle est maximale (*apogée*), 407 000 kilomètres.

Comme les Grecs l'avaient pensé, la Lune est de loin le corps céleste le plus proche de la Terre. Sans parler des étoiles, et en nous limitant au système solaire, la Lune se trouve, en proportion, à notre porte.

Son diamètre, estimé à partir de sa distance et de sa taille apparente, vaut 3 480 kilomètres. Celui de la Terre vaut 3,65 fois plus, et celui du Soleil 412 fois plus. Il se trouve que la distance Soleil-Terre vaut environ 390 fois la distance moyenne Terre-Lune, ce qui fait que les différences entre les tailles et les distances se compensent presque exactement, et que les deux corps, de tailles si différentes en réalité, nous *apparaissent* presque de même taille dans le ciel. C'est pourquoi, quand la Lune passe juste devant le Soleil, ce corps beaucoup plus petit, mais beaucoup plus proche, masque presque parfaitement le Soleil, plus grand mais plus éloigné, ce qui fait d'une éclipse totale du Soleil un merveilleux spectacle. C'est là une coïncidence tout à fait étonnante – et très agréable.

ALLER DANS LA LUNE

La proximité relative de la Lune et sa place d'honneur dans le ciel stimulent depuis longtemps l'imagination des hommes : y a-t-il un moyen de l'atteindre ? (On aurait pu aussi rêver d'un moyen d'atteindre le Soleil, mais sa chaleur intense ne pouvait pas manquer de refroidir l'imagination... De toute évidence, la Lune constituait un objectif à la fois moins rébarbatif et plus proche.)

Aux temps primitifs, cela ne semblait pas une tâche insurmontable d'atteindre la Lune : on supposait que l'atmosphère s'étendait jusqu'aux corps célestes, de sorte qu'un système capable de soulever un homme dans les airs pouvait aussi bien, à la limite, l'emmener jusqu'à la Lune.

C'est ainsi qu'au IIe siècle apr. J.-C., l'écrivain syrien Lucien de Samosate publia la première histoire de voyage spatial que nous connaissions. Dans cette histoire, un navire est pris dans une sorte de jet d'eau qui le soulève très haut, assez haut pour atteindre la Lune.

En 1638 parut *L'Homme dans la Lune,* écrit par un prêtre anglais, Francis Godwin (qui mourut avant la publication). Le héros de Godwin est entraîné jusqu'à la Lune dans un chariot remorqué par de grandes oies, que leur migration annuelle emmène jusque là-bas.

En 1643, toutefois, on commença à comprendre la nature de la pression atmosphérique, et on se rendit rapidement compte que l'atmosphère ne pouvait pas s'étendre très loin au-dessus de la surface terrestre. La plus grande partie de l'espace entre la Terre et la Lune était du vide, dans lequel les jets d'eau ne pouvaient pas monter et où les oies ne pouvaient pas voler. Le voyage dans la Lune paraissait soudain beaucoup plus difficile, mais pourtant pas impossible.

Quelques années plus tard, ce fut la parution du *Voyage dans la Lune* de Cyrano de Bergerac, écrivain et duelliste français. Cyrano y évoque sept moyens d'atteindre la Lune. Six de ces moyens sont irréalisables pour une raison ou pour une autre, mais le septième repose sur l'emploi des fusées. Celles-ci étaient en effet à l'époque (et sont encore, d'ailleurs) le seul moyen de propulsion utilisable dans le vide.

Mais ce n'est pas avant 1687 que l'on comprit le principe des fusées. Cette année-là, Newton publia ses *Principia Mathematica* où il énonça, entre autres choses, ses trois lois du mouvement. La troisième est généralement connue sous le nom de « loi de l'action et de la réaction » : quand une force s'exerce dans un sens, une force opposée s'exerce dans l'autre. Quand une fusée éjecte une masse de matière dans un sens, le reste de la fusée se déplace dans l'autre sens, aussi bien dans le vide que dans l'air. Mieux, même, car dans le vide il n'y a pas de résistance opposée par l'air au déplacement. C'est à tort qu'on imagine généralement qu'une fusée a besoin de « pousser contre quelque chose ».

LES FUSÉES

D'ailleurs, les fusées n'étaient pas seulement un sujet de réflexion théorique : elles existaient des siècles avant que Cyrano écrive et que Newton élabore sa théorie.

Dès le XIIIe siècle, les Chinois inventèrent et utilisèrent de petites fusées, qui visaient seulement un objectif psychologique : effrayer l'ennemi. C'est la civilisation occidentale moderne qui allait les adapter à des usages plus

sanguinaires. En 1801, un expert anglais en matière d'artillerie, William Congreve, apprit en Orient l'existence des fusées que les troupes indiennes utilisaient contre les Anglais dans les années 1780. Il mit au point un certain nombre de missiles redoutables, dont certains furent utilisés contre les États-Unis dans la guerre de 1812, en particulier lors du bombardement de Fort McHenry en 1814. Les progrès de l'artillerie conventionnelle en portée, en précision et en puissance firent oublier les fusées. Mais la Seconde Guerre mondiale vit se développer le « bazooka » américain et la « katioucha » russe, qui sont tous les deux des paquets d'explosif propulsés par des fusées. Les avions à réaction, à une échelle beaucoup plus grande, utilisent aussi le principe d'action et de réaction.

Vers le début du XXe siècle, deux hommes imaginèrent, indépendamment l'un de l'autre, un usage nouveau, et plus noble, des fusées : l'exploration de la haute atmosphère et de l'espace. L'un était russe, Konstantine Edouardovitch Tsiolkovski, et l'autre américain, Robert Hutchings Goddard. (Il est curieux, compte tenu de la suite de l'histoire, qu'un Russe et un Américain aient été les pionniers de l'âge des fusées, encore qu'un inventeur allemand fort imaginatif, Hermann Ganswindt, ait publié à la même époque des spéculations encore plus hardies, bien que moins systématiques et moins scientifiques.)

Le Russe a publié le premier ses réflexions et ses calculs, de 1903 à 1913, tandis que Goddard n'a publié qu'à partir de 1919. Mais Goddard a été le premier à mettre ses idées en pratique. Le 16 mars 1926, d'une ferme enneigée d'Auburn, dans le Massachusetts, il a lancé une fusée qui s'est élevée à soixante mètres. Chose remarquable, cette fusée utilisait un combustible liquide au lieu de poudre à canon. D'autre part, contrairement aux fusées ordinaires, bazookas, avions à réaction, etc... qui brûlent grâce à l'oxygène de l'air environnant, la fusée de Goddard, destinée à fonctionner dans le vide spatial, devait transporter son propre comburant sous forme d'oxygène liquide (*lox* dans l'argot moderne des spécialistes).

Jules Verne, dans ses romans de science-fiction du XIXe siècle, utilisait un canon pour envoyer un projectile vers la Lune. Mais le canon fournit toute son énergie d'un seul coup, au départ, au moment où l'atmosphère est la plus dense, et oppose la plus forte résistance. De plus, l'accélération nécessaire pour obtenir dès le départ la vitesse maximale est tellement forte qu'elle réduirait en bouillie des passagers humains.

Au contraire, les fusées de Goddard montaient doucement au début, accéléraient et dépensaient leur dernière poussée très haut dans l'atmosphère la plus ténue, où la résistance est faible. L'augmentation progressive de la vitesse permettait à l'accélération de ne pas dépasser un niveau supportable, ce qui est évidemment un point important pour des véhicules habités.

Malheureusement, le succès de Goddard passa à peu près inaperçu, sauf de ses voisins indignés qui l'obligèrent à aller poursuivre ses essais ailleurs. Goddard s'isola donc pour lancer ses fusées. Entre 1930 et 1935, ses véhicules atteignaient des vitesses proches de 1 000 km/h, et des altitudes de près de trois kilomètres. Il mit au point des systèmes de guidage en vol et des stabilisateurs gyroscopiques pour garder la fusée pointée dans la bonne direction. Il fit breveter l'idée de fusées à plusieurs étages. En effet, ce dispositif permet de se débarrasser en plusieurs fois du poids mort, et d'atteindre des vitesses et des altitudes beaucoup plus grandes qu'avec la même quantité totale de carburant dans une fusée à un seul étage.

Pendant la Seconde Guerre mondiale, la marine américaine finança sans conviction de nouvelles expériences de Goddard. Pendant ce temps, le gouvernement allemand investissait dans ce domaine un effort considérable. Le groupe de jeunes chercheurs qui s'y consacra était inspiré principalement par Hermann Oberth, un mathématicien roumain qui, indépendamment de Tsiolkovski et Goddard, avait publié des études sur les fusées et les voyages dans l'espace dès 1923. Les recherches allemandes commencèrent en 1935 et culminèrent avec la mise au point du V2. Sous la direction de l'expert en fusées Wernher von Braun (qui proposa ses services aux États-Unis à la fin de la guerre), le premier missile propulsé par une fusée était lancé en 1942. Le V2 entrait en service en 1944, trop tard pour permettre aux Nazis de gagner la guerre, bien qu'ils aient eu le temps d'en lancer en tout 4 300, dont 1 230 atteignirent Londres. Les missiles de von Braun tuèrent ainsi 2 511 Anglais et en blessèrent grièvement 5 869 autres.

Goddard est mort le 10 août 1945, presque le dernier jour de la guerre. Il a tout juste eu le temps de voir enfin s'épanouir en flamme l'étincelle qu'il avait lancée. Stimulés par les succès du V2, les États-Unis et l'Union Soviétique se sont lancés à fond dans la recherche sur les fusées, chacun s'assurant le concours des experts allemands sur lesquels il pouvait mettre la main.

Les États-Unis ont commencé par utiliser des V2 pris aux Allemands pour étudier la haute atmosphère, mais le stock s'est épuisé en 1952. A ce moment-là, on fabriquait déjà des fusées plus grandes et plus perfectionnées, aussi bien aux États-Unis qu'en Union Soviétique, et les progrès ont continué.

EXPLORER LA LUNE

Une ère nouvelle a commencé le 4 octobre 1957 (à un mois du centenaire de la naissance de Tsiolkovski), quand l'Union Soviétique a mis en orbite le premier satellite artificiel (*Spoutnik I*). Celui-ci a parcouru autour de la Terre une orbite elliptique, et il est passé à 252 kilomètres de la surface au périgée, à 900 kilomètres à l'apogée. Une orbite elliptique ressemble au « grand huit » des attractions foraines. En allant de l'apogée au périgée, le satellite descend et perd de l'énergie potentielle : sa vitesse augmente, de sorte qu'au périgée il commence à remonter avec une vitesse maximale, comme le chariot du grand huit. Sa vitesse diminue à mesure qu'il grimpe (comme celle du chariot) et passe par un minimum au moment où, ayant atteint l'apogée, il commence à descendre à nouveau.

Au périgée, *Spoutnik I* traversait les régions ténues de la haute atmosphère, et la résistance de l'air, bien que minime, le ralentissait un peu à chaque passage. A chaque tour, l'apogée était un peu moins haut qu'au tour précédent. Le satellite décrivait lentement une spirale descendante. Il devait finir par plonger dans les régions plus denses de l'atmosphère, et y brûler par suite du frottement de l'air.

La vitesse à laquelle l'orbite d'un satellite change, pour la raison que nous venons d'évoquer, dépend en partie de la masse du satellite, en partie de sa forme et en partie de la densité de l'air qu'il traverse. Aussi, cela permet de calculer à chaque altitude la densité de l'atmosphère. Pour la haute atmosphère, ce sont les satellites qui nous ont donné les premières

mesures de densité. Celle-ci s'est révélée un peu plus forte que prévu ; mais elle ne vaut quand même, à une altitude de 240 kilomètres que 1/10 000 000 de sa valeur au niveau de la mer, et 1/1 000 000 000 à 330 kilomètres.

Ces souffles impalpables ne sont pas négligeables pour autant. Même à une hauteur de 1 600 kilomètres où la densité atmosphérique est un million de millions de fois plus faible qu'au niveau de la mer, elle est tout de même un milliard de fois plus forte que celle du gaz qui occupe le « vide » extérieur. L'enveloppe gazeuse de la Terre s'étend très loin.

L'Union Soviétique n'est pas longtemps restée seule dans l'espace : quatre mois plus tard, le 30 janvier 1958, les États-Unis mettaient en orbite leur premier satellite, *Explorer I.*

Une fois des satellites mis en orbite autour de la Terre, les yeux se sont tournés avec un intérêt accru vers la Lune. Bien sûr, celle-ci avait perdu dans l'intervalle une partie de son charme : bien que ce fût un monde, et non pas une simple lumière dans le ciel, ce n'était plus le monde dont on avait d'abord rêvé.

Avant les observations de Galilée, on avait toujours supposé que si les corps célestes étaient des mondes, ils portaient forcément des êtres vivants, voire des êtres intelligents, ressemblant aux hommes. Les premières histoires de science-fiction à propos de la Lune partaient de cette hypothèse, et les suivantes aussi, au début du XXe siècle.

En 1835, un écrivain anglais nommé Richard Adams Locke publia dans le *New York Sun* une série d'articles prétendant décrire des études scientifiques sérieuses de la surface lunaire, études qui auraient permis d'y découvrir de nombreuses espèces vivantes. Les descriptions étaient détaillées, et des millions de personnes se laissèrent tromper.

Pourtant, après que Galilée eut publié ses premières observations de la Lune à travers la lunette, il n'avait pas fallu longtemps pour que l'on commence à se rendre compte qu'il ne pouvait exister aucune vie sur la Lune. Sa surface n'était jamais cachée par des nuages ni du brouillard. La limite entre la partie éclairée et la partie sombre était toujours parfaitement nette, ce qui prouvait l'absence de crépuscule. Les « mers » sombres, que Galilée avait prises pour des étendues d'eau, s'étaient vite révélées criblées de petits cratères : ce n'était donc que des plaines relativement lisses, du sable sans doute. La Lune ne possédait ni eau ni air : elle ne pouvait pas être habitée.

C'était peut-être là une conclusion néanmoins un peu rapide : et la face cachée, celle que les hommes ne voyaient jamais ? Et ne pouvait-il pas y avoir un peu d'eau sous la surface, en quantité insuffisante pour des animaux de grande taille, mais suffisante peut-être pour permettre l'existence d'êtres analogues aux bactéries ? Et même si aucune vie n'était présente, ne pouvait-on pas trouver dans le sol des corps chimiques représentant une étape – peut-être avortée – d'une lente évolution vers la vie ? Même s'il n'y avait rien de tout cela, la Lune posait beaucoup de questions qui n'avaient rien à voir avec la vie. Où s'était-elle formée ? Quelle était sa composition minéralogique ? Quel âge avait-elle ?

Aussi, c'est très vite après le lancement de *Spoutnik I* que les nouvelles techniques se tournèrent vers la Lune. Dès le 2 janvier 1959, l'Union Soviétique lançait la première *sonde lunaire,* c'est-à-dire le premier satellite à passer près de la Lune. C'était *Lunik I,* le premier engin placé sur orbite autour du Soleil. Deux mois plus tard, les États-Unis en faisaient autant.

Le 12 septembre 1959, les Russes lançaient *Lunik II*, pointé de manière à heurter la surface lunaire. Pour la première fois dans l'Histoire, un objet fabriqué par l'homme reposait sur la surface d'un autre monde. Un mois plus tard, le véhicule soviétique *Lunik III* passait derrière la Lune et pointait une caméra de télévision vers la face cachée. Quarante minutes d'images de cette face atteignirent la Terre. Elles étaient prises d'une distance de 60 000 kilomètres. Elles étaient assez floues, mais révélaient tout de même quelque chose d'intéressant : l'autre face de la Lune était pratiquement dépourvue des « mers » qui sont un des aspects les plus remarquables de la face tournée vers la Terre. Pourquoi cette dissymétrie ? Peut-être les mers s'étaient-elles formées assez tard dans l'histoire de la Lune, quand une face était déjà tournée en permanence vers la Terre, dont l'attraction déviait vers cette face les grandes météorites qui sont à l'origine de la formation des « mers » ?

Mais l'exploration lunaire ne faisait que commencer. En 1964, les États-Unis lançaient la sonde *Ranger 7*, qui devait s'écraser sur la surface de la Lune, en prenant des photos jusqu'au dernier moment. Le 31 juillet 1964, on obtenait ainsi 4 316 photos d'une région que l'on appelle maintenant *mare cognitum* (la mer connue). Début 1965, *Ranger 8* et *Ranger 9* montraient que la surface de la Lune était dure, au plus friable, au lieu d'être couverte d'une épaisse couche de poussières comme le pensaient de nombreux astronomes. Les sondes révélèrent que même les régions qui semblaient les plus plates, vues au télescope, étaient en réalité criblées de cratères trop petits pour être visibles depuis la Terre.

Le 3 février 1966, la sonde soviétique *Luna IX* réussissait un *alunissage en douceur* – c'est-à-dire qui n'entraîne pas la destruction de l'engin – et envoyait des photos prises au niveau du sol lunaire. Deux mois plus tard, les Russes mettaient *Luna X* en orbite autour de la Lune, orbite que la sonde parcourait en trois heures. Cette sonde a mesuré la radioactivité de la surface lunaire, et ses résultats ont permis d'estimer que les roches de la surface étaient analogues au basalte qui couvre le fond des océans terrestres.

Les Américains ont emboîté le pas, avec des fusées plus perfectionnées encore. *Surveyor 1* a « aluni » en douceur le 1er juin 1966, et en septembre 1967, *Surveyor 5* était capable de manipuler et d'analyser des échantillons du sol, sous télécommande radio depuis la Terre : le sol était effectivement basaltique, et contenait des particules de fer qui avaient probablement une origine météoritique.

Le 10 août 1966, la première sonde américaine de la série Lunar Orbiter était mise sur orbite autour de la Lune. Ces sondes ont photographié en détail toute la Lune, de sorte que toute la surface – y compris la partie toujours invisible de la Terre – est parfaitement connue. De plus, ces sondes ont pris des photos saisissantes de la Terre, vue du voisinage de la Lune.

Les cratères ont reçu les noms de certains astronomes et autres grands personnages du passé. Comme la plupart ont été choisis par l'astronome italien Giovanni Battista Riccioli vers 1650, les plus grands cratères immortalisent les astronomes les plus anciens (Copernic, Tycho, Kepler) ou les Grecs de l'Antiquité (Aristote, Archimède, Ptolémée).

L'autre face, révélée pour la première fois par *Lunik III*, a permis une seconde série de baptêmes. Bien entendu, les Russes (premier arrivé,

premier servi...) se sont réservé les sites les plus remarquables. Ils ont dédié ainsi des cratères non seulement à Tsiolkovsky, le grand prophète du voyage spatial, mais aussi à Lomonosov et Popov, deux chimistes russes de la fin du XVIII[e] siècle. Ils ont honoré de la même manière un certain nombre de personnalités occidentales, en particulier Maxwell, Hertz, Edison, Pasteur et les Curie, qui sont tous mentionnés dans ce livre. Un nom bien à sa place de l'autre côté de la Lune : celui du pionnier français de la science-fiction, Jules Verne.

En 1970, la face cachée était suffisamment connue pour que l'on puisse baptiser systématiquement tous les sites dignes d'intérêt. Sous la direction de l'astronome américain Donald Howard Menzel, une commission internationale a procédé à des centaines de baptêmes, honorant ainsi des hommes ayant contribué aux progrès de la science dans un domaine ou dans un autre. De beaux cratères sont ainsi dédiés aux Russes Mendeleïev (auteur de la première table périodique des éléments, dont nous parlerons au chapitre 6) et Gagarine, premier homme placé en orbite autour de la Terre, et mort depuis dans un accident d'avion. D'autres sites honorent l'astronome hollandais Hertzsprung, le mathématicien français Galois, le physicien italien Fermi, le mathématicien américain Wiener et le physicien anglais Cockcroft. Dans un secteur bien délimité, on rencontre Nernst, Roentgen, Lorentz, Moseley, Einstein, Bohr et Dalton, en l'honneur du rôle important qu'ils ont tous joué dans l'élaboration de la théorie atomique et l'étude de la structure de l'atome.

L'intérêt de Menzel pour les ouvrages scientifiques et la science-fiction se révèle dans sa juste décision d'attribuer quelques cratères à ceux qui ont su éveiller l'enthousiasme de toute une génération pour les voyages spatiaux, à un moment où la science orthodoxe les considérait comme une chimère. Un cratère est ainsi dédié à Hugo Gernsback, éditeur des premiers magazines américains entièrement consacrés à la science-fiction, et un autre à Willy Ley qui est, de tous les écrivains, le plus ardent prophète des victoires et des possibilités de la fusée.

LES ASTRONAUTES ET LA LUNE

Mais si passionnante et fructueuse fût-elle, l'exploration de la Lune par les automates ne suffisait pas. N'était-il donc pas possible que les fusées emmènent des hommes ? De fait, le premier pas dans cette direction a été fait trois ans et demi à peine après le lancement de *Spoutnik I.*

Le 12 avril 1961, le cosmonaute soviétique Youri Alexeïevitch Gagarine était placé sur orbite et revenait sain et sauf sur Terre. Trois mois plus tard, un autre cosmonaute soviétique, Gherman Stepanovitch Titov, décrivait dix-sept orbites avant d'atterrir, passant ainsi vingt-quatre heures en apesanteur. Le 20 février 1962, c'était le tour de l'Américain John Herschel Glenn, qui parcourut trois fois l'orbite de la Terre. Depuis lors, des douzaines d'hommes ont quitté la Terre, et certains sont restés des mois dans l'espace. Le 16 juin 1963, une cosmonaute soviétique, Valentina Tereshkova, partait pour un vol de soixante et onze heures, décrivant en tout dix-sept tours autour de la Terre. Tout juste vingt ans plus tard, l'Américaine Sally Ride la suivait dans l'espace.

Les fusées ont emmené bientôt des équipages de plusieurs hommes : les Soviétiques Vladimir Komarov, Konstantin Feokstitov et Boris Yegorov le 12 octobre 1964, les Américains Virgil Grissom et John Young le 23 mars 1965.

Les cosmonautes ont quitté leur véhicule dans l'espace : Aleksei Leonov (URSS) le 18 mars 1965 et Edward White (USA) le 3 juin 1965.

Jusque-là, la plupart des « premières » de l'espace étaient à l'actif des Soviétiques. Mais à partir de cette date, les Américains allaient prendre la tête de la « course à l'espace ». Les véhicules manœuvraient dans l'espace, s'y rejoignaient, s'y accrochaient l'un à l'autre et commençaient à s'éloigner de plus en plus de la Terre.

Ce programme, hélas, ne s'est pas déroulé sans deuils. En janvier 1967, trois cosmonautes américains (Grissom, White et Chaffee) sont morts dans l'incendie de leur capsule, au cours d'essais au sol. Le 23 avril 1967, Komarov était victime d'un mauvais fonctionnement du parachute d'atterrissage : il a été le premier à mourir au cours d'une mission spatiale.

Les plans américains de missions de trois hommes vers la Lune (programme Apollo) ont été retardés à la suite de ce tragique accident : il fallait revoir la sécurité de la capsule. Mais ils n'ont pas été abandonnés. Le premier véhicule Apollo à emmener un équipage, *Apollo 7*, a décollé le 21 octobre 1968, sous le commandement de Walter Schirra. Le 20 décembre, sous le commandement de Frank Borman, *Apollo 8* a fait le tour de la Lune à faible distance. *Apollo 10*, lancé le 18 mai 1969, s'est approché de la Lune et a largué son module lunaire, qui est descendu à quinze kilomètres de la surface.

Enfin, le 16 juillet 1969, *Apollo 11* a décollé sous le commandement de Neil Armstrong. Le 20 juillet, celui-ci était le premier homme à marcher sur la Lune.

Depuis lors, six autres Apollo ont pris le départ. Cinq de ces missions ont été couronnées de succès. Seul *Apollo 13* a dû faire demi-tour sans remplir sa mission – mais en sauvant l'équipage !

Le programme soviétique ne comporte pas de vols habités vers la Lune. Mais le 12 septembre 1970, un véhicule automatique russe a aluni en douceur, a recueilli des échantillons du sol et les a rapportés sur Terre. Plus tard, un autre automate s'est promené sur la surface lunaire pendant plusieurs mois, sous contrôle radio, et a envoyé vers la Terre toutes sortes de mesures et de renseignements.

L'étude des pierres lunaires rapportées sur Terre par tous ces vols, automatiques ou humains, montre que la Lune est un monde absolument mort. Sa surface semble avoir été soumise à des températures très élevées, car elle est couverte de fragments vitreux, ce qui indiquerait que les roches superficielles ont fondu. Aucune trace d'eau n'a pu être détectée, ni aucun indice que l'eau puisse être présente en profondeur, ou même avoir été présente dans le passé. Il n'y a aucune vie, aucune trace de matériaux chimiques ayant quoi que ce soit à voir avec la matière vivante.

Depuis 1971, aucun véhicule n'a atterri sur la Lune et aucune expédition n'est actuellement prévue. La preuve est faite, néanmoins, que la technologie humaine est capable de transporter des hommes et des machines sur la Lune dès que cela semblera souhaitable. Pour l'instant, les programmes spatiaux se poursuivent dans d'autres directions.

Vénus et Mercure

Des planètes qui tournent autour du Soleil, deux seulement – Vénus et Mercure – en sont plus proches que la Terre. Alors que la Terre est, en moyenne, à 150 millions de kilomètres du Soleil, Vénus en est à 108 millions de kilomètres, et Mercure à 58 millions de kilomètres.

Par conséquent, Vénus et Mercure nous apparaissent toujours relativement près du Soleil. Vénus ne s'en écarte jamais, pour nous, de plus de 47°, et Mercure de 28°. Quand elle est à l'est du Soleil, Vénus apparaît dans le ciel de l'Ouest juste après le coucher du Soleil – et on l'appelle alors *étoile du soir*. Quand elle est de l'autre côté de son orbite, à l'ouest du Soleil, elle se montre juste avant l'aube, se levant à l'est un peu avant le Soleil – et on l'appelle *étoile du matin*.

Il a d'abord semblé naturel de penser que l'« étoile » du soir et celle du matin devaient être deux corps célestes différents. Mais quand l'une est dans le ciel, l'autre n'y est pas. Il est alors devenu peu à peu évident qu'il s'agissait du même corps, qui se montrait tantôt d'un côté du Soleil, tantôt de l'autre, jouant ainsi alternativement le rôle d'étoile du matin et d'étoile du soir. Le premier Grec à exprimer cette idée fut Pythagore, au VIe siècle av. J.-C., mais elle lui venait peut-être des Babyloniens.

Des deux planètes, Vénus est de loin la plus facile à observer. D'abord, elle est plus près de la Terre : quand elles sont toutes les deux du même côté du Soleil, leur distance peut atteindre le minimum de 40 millions de kilomètres. Vénus est alors seulement cent fois plus loin de nous que la Lune. Aucun corps céleste important (à part la Lune) ne nous approche d'aussi près que Vénus. En revanche, la distance moyenne entre Mercure et la Terre, quand les deux planètes sont du même côté du Soleil, vaut 90 millions de kilomètres.

Non seulement Vénus est plus proche de la Terre (au moins quand elles sont bien placées toutes les deux), mais elle est plus grosse que Mercure et renvoie plus de lumière. Elle a en effet un diamètre de 12 000 kilomètres, tandis que celui de Mercure ne vaut que 4 800 kilomètres. Enfin, Vénus est couverte de nuages et renvoie une fraction bien plus importante de la lumière qu'elle reçoit du Soleil que ne peut le faire Mercure, qui n'a pas d'atmosphère et ne dispose pour renvoyer la lumière que de roche nue (comme la Lune).

Aussi, Vénus a une magnitude de – 4,22 (quand elle brille le plus). Elle est alors 12,6 fois plus brillante que Sirius, la plus brillante des étoiles. En fait, c'est l'objet le plus brillant du ciel après le Soleil et la Lune : par les nuits sans lune, Vénus projette parfois des ombres perceptibles. Mercure, à son maximum, n'a qu'une magnitude de – 1,2, ce qui en fait un objet presque aussi brillant que Sirius, mais dix-sept fois moins que Vénus (à son maximum).

De plus, Mercure est si proche du Soleil qu'on ne la voit que près de l'horizon, et à des moments où l'aube – ou le crépuscule – illumine le ciel. Aussi, malgré son éclat, c'est une planète difficile à observer. Il semble, par exemple, que Copernic ne l'ait jamais vue.

Ces deux planètes étant proches du Soleil et se montrant tantôt d'un côté, tantôt de l'autre, il était inévitable que certains imaginent qu'elles tournaient autour du Soleil plutôt qu'autour de la Terre. Suggérée par plusieurs anciens Grecs, cette idée sera pourtant rejetée, jusqu'à ce qu'elle

soit reprise 1 900 ans plus tard par Tycho, puis que Copernic l'étende à toutes les planètes.

Si Copernic avait raison et si Vénus était un corps opaque, ne brillant (comme la Lune) qu'en renvoyant la lumière reçue du Soleil, alors, vue de la Terre, elle devait avoir des phases, tout comme la Lune. C'est ce que constata Galilée, à partir du 11 décembre 1610, grâce à sa lunette. Des observations répétées lui en ont donné bientôt la certitude : Vénus avait des phases comme la Lune. C'était le coup de grâce pour le vieux système du monde, avec la Terre au centre : il ne pouvait pas rendre compte des phases de Vénus, telles que l'on pouvait les observer. Mercure aussi, plus tard, se révéla avoir des phases.

MESURER LES PLANÈTES

Les deux planètes étaient difficiles à observer au télescope. Mercure était si près du Soleil, si petite et si éloignée qu'on ne pouvait pas distinguer beaucoup de détails à sa surface. Pourtant, l'astronome italien Giovanni Virginio Schiaparelli étudia ces détails soigneusement d'une façon régulière, et à partir de leur changement avec le temps, annonça en 1889 que Mercure tournait sur elle-même en 88 jours.

Ce résultat semblait raisonnable, car Mercure tourne autour du Soleil en 88 jours également. Elle est assez près du Soleil pour lui être « fixée » par la gravitation, comme la Lune l'est à la Terre, et lui montrer toujours la même face. Dans ce cas, sa période de rotation (sur son axe) devrait être égale à sa période de révolution (autour du Soleil).

Vénus, bien que plus large et plus proche, est encore plus difficile à observer au télescope, parce qu'elle est perpétuellement couverte d'une épaisse couche de nuages sans aucune éclaircie, et ne révèle à l'observatoire qu'un manteau blanc absolument uniforme. Aussi, ces observations ne permettent absolument pas de déterminer sa période de rotation. Faute de mieux, certains astronomes supposaient qu'elle aussi devait être « verrouillée » au Soleil par la gravitation, avec une période de rotation égale à sa période de révolution (224,7 jours).

Telle était la situation jusqu'à la mise au point des techniques radar, qui ont permis d'émettre un faisceau d'ondes centimétriques et de détecter le faisceau renvoyé par un obstacle. Mis au point pendant la Seconde Guerre mondiale pour détecter les avions, le radar pouvait aussi « rebondir » sur des objets célestes.

Dès 1946, le Hongrois Zoltan Lajos Bay envoyait un faisceau vers la Lune et détectait l'écho qu'elle lui renvoyait.

Mais la Lune est, en comparaison, une cible facile. En 1961, trois équipes américaines, une anglaise et une soviétique ont réussi à envoyer des faisceaux vers Vénus et à détecter les échos qu'elle renvoyait. Ces faisceaux se déplaçaient à la vitesse de la lumière, qui était alors connue avec précision. En mesurant le temps mis par le faisceau pour faire l'aller-retour Terre-Vénus, on a pu calculer la distance Terre-Vénus (à ce moment-là) avec une précision beaucoup plus grande que par les méthodes utilisées auparavant. Les proportions du système solaire étant parfaitement connues, on a pu en déduire, avec la même précision, les distances entre deux quelconques des objets qui le constituent.

De plus, tout objet qui n'est pas au zéro absolu (et aucun objet ne peut y être...) émet un rayonnement qui peut être dans le domaine visible, infrarouge ou dans celui des ondes millimétriques. A partir des longueurs d'onde de ce rayonnement, on peut connaître la température de l'objet.

En 1962, on est parvenu à détecter le rayonnement émis par la partie sombre de Mercure, celle qui n'est pas éclairée par le Soleil. Si la période de rotation de Mercure est bien de 88 jours, c'est toujours la même face qui est tournée vers le Soleil : elle doit être très chaude, et l'autre très froide. Or, l'étude du rayonnement détecté à partir de cette face sombre révélait une température beaucoup moins basse que prévu, ce qui prouvait qu'elle n'était pas toujours dans la nuit. Quelle était donc la période de rotation de Mercure ?

C'est encore le radar qui allait fournir la réponse. Quand le faisceau « rebondit » sur un corps en mouvement, le faisceau qui revient est légèrement modifié, d'une façon qui permet de calculer la vitesse de l'objet qui l'a renvoyé. En 1965, deux Américains, Rolf Buchanan Dyce et Gordon H. Pettengill, se sont aperçus ainsi que la surface de Mercure tournait plus vite que prévu : Mercure tournait autour de son axe en 59 jours, ce qui faisait que chaque portion de la surface se trouvait de temps en temps au Soleil.

La période exacte de rotation (58,65 jours) vaut exactement les deux tiers de la période de révolution (88 jours). C'est une autre forme de « verrouillage » gravitationnel, mais moins serré que celui où les deux périodes sont égales.

LES SONDES EXPLORENT VÉNUS

Vénus réservait bien d'autres surprises. Comme elle a à peu près la même taille que la Terre (12 000 kilomètres de diamètre au lieu de 12 800), on y voyait souvent une sorte de « sœur jumelle » de la Terre. Bien sûr, elle est plus proche du Soleil, mais on pensait que son manteau de nuages pouvait l'empêcher de devenir trop chaude. On supposait ces nuages formés de gouttelettes d'eau, et on en concluait que Vénus devait avoir des océans, plus vastes peut-être que ceux de la Terre, et qu'elle pouvait abriter une vie marine foisonnante. Beaucoup d'histoires de science-fiction (y compris plusieurs des miennes) décrivent cette planète riche en eau et grouillante de vie.

Le premier choc est arrivé en 1956. Une équipe américaine, dirigée par Cornell H. Mayer, qui étudiait le rayonnement émis par la face sombre de Vénus, en a conclu que cette face était à une température bien supérieure au point d'ébullition de l'eau.

Cette conclusion était presque incroyable. On avait besoin de quelque chose de plus convaincant qu'un petit faisceau d'ondes. Une fois les fusées suffisamment au point pour fréquenter les parages de la Lune, il semblait logique d'envoyer des sondes vers d'autres planètes.

Le 27 août 1962, les États-Unis lançaient la première sonde qui devait réussir à approcher Vénus, *Mariner 2*. Elle emporta des instruments destinés à détecter et analyser le rayonnement émis par Vénus, puis à transmettre les résultats à la Terre, à travers des dizaines de millions de kilomètres de vide.

Le 14 décembre 1962, *Mariner 2* passait à 32 000 kilomètres au-dessus des nuages de Vénus, et il ne restait plus de doute : la planète était un enfer brûlant, aux pôles comme à l'équateur, sur la face sombre comme sur la face éclairée. La température de surface est voisine de 475° C, plus qu'il n'en faut pour fondre l'étain et le plomb, et faire bouillir le mercure.

Ce ne fut pas le seul résultat pour l'année 1962. Les faisceaux radar traversent les nuages. Les faisceaux envoyés vers Vénus traversaient la couche de nuages et « rebondissaient » sur le sol : ils pouvaient « voir » la surface, contrairement aux yeux de l'homme, qui ne détectent que la lumière « visible ». En 1962, étudiant les modifications subies par le faisceau réfléchi, Roland L. Carpenter et Richard M. Goldstein ont constaté que Vénus avait une période de rotation de l'ordre de 250 jours terrestres. Une analyse ultérieure par le physicien américain Irwin Ira Shapiro a précisé cette période : 243,09 jours. Cette lenteur ne résulte pas d'un « verrouillage » sur le Soleil, puisque la période de révolution est de 224,7 jours. Vénus tourne donc *plus lentement* sur elle-même qu'elle ne tourne autour du Soleil.

De plus, elle tourne sur elle-même « dans le mauvais sens ». Vue d'un point situé dans la direction de l'Étoile polaire, Vénus tourne sur elle-même dans le sens des aiguilles d'une montre, alors que la quasi-totalité des mouvements de rotation et de révolution, dans le système solaire, se font dans l'autre sens. On n'a pas encore trouvé d'explication satisfaisante de ce mouvement *rétrograde*.

Chaque fois que Vénus est à son point le plus proche de la Terre, elle nous présente la même face, car elle a tourné exactement cinq fois sur elle-même (dans le mauvais sens) depuis la fois précédente. S'agit-il d'un « verrouillage » sur la Terre ? Celle-ci semble trop petite pour cela, compte tenu de la distance qui sépare les deux planètes.

Après *Mariner 2,* les sondes se sont succédé, lancées aussi bien par les États-Unis que par l'Union Soviétique. Celles de l'URSS étaient conçues pour pénétrer dans l'atmosphère de Vénus, puis y descendre en parachute jusqu'à se poser en douceur. Mais les conditions sur Vénus sont si dures que ces sondes cessaient de fonctionner très vite après leur arrivée au sol, non sans avoir envoyé auparavant des renseignements précieux sur l'atmosphère vénusienne.

D'abord, cette atmosphère est étonnamment dense (90 fois plus que celle de la Terre) et se compose surtout de gaz carbonique (qui n'intervient que pour une faible part dans la composition de l'atmosphère terrestre). L'atmosphère de Vénus se compose de 96,6 % de gaz carbonique et 3,2 % d'azote (mais la densité de l'atmosphère étant ce qu'elle est, cette atmosphère contient ainsi trois fois plus d'azote que celle de la Terre !).

Le 20 mai 1978, les États-Unis lançaient *Pioneer Venus,* qui atteignit la planète le 4 décembre, et se plaça en orbite autour d'elle. Cette station spatiale passa pratiquement au-dessus des pôles de Vénus. Elle lâcha plusieurs sondes, qui entrèrent dans l'atmosphère, confirmant et précisant les résultats déjà obtenus par les Russes.

La principale couche de nuages a un peu plus de trois kilomètres d'épaisseur et se trouve à près de cinquante kilomètres de la surface. Cette couche est formée d'eau, contenant une certaine quantité de soufre. Elle est surmontée d'une brume d'acide sulfurique corrosif.

Sous la couche de nuages s'étend une zone brumeuse de quinze kilomètres d'épaisseur, et plus bas l'atmosphère est complètement transparente. Dans cette région basse, l'atmosphère semble parfaitement calme, sans tempêtes et sans changements de « temps » – seulement une chaleur intense et constante. Il n'y a que des vents très faibles, mais compte tenu de la densité du gaz, un souffle de brise doit avoir la même force qu'un ouragan sur la Terre. Finalement, on imagine difficilement un monde plus désagréable que la « sœur jumelle de la Terre ».

Le rayonnement solaire qui atteint Vénus est presque complètement renvoyé ou absorbé par les nuages, et seulement 3 % de ce rayonnement parvient dans la partie claire de l'atmosphère. Le sol, lui, doit en recevoir environ 2,5 %. Étant plus proche du Soleil que la Terre, et recevant donc au départ une énergie plus grande, Vénus ne reçoit à la surface du sol que le sixième de l'énergie qui atteint le sol terrestre, à cause de son épaisse couche de nuages. Aussi la lumière y est-elle plus faible que sur la Terre. Toutefois, si nous étions capables de vivre sur Vénus, nous y verrions sans peine.

D'ailleurs, l'une des sondes soviétiques a eu le temps, après l'atterrissage sur la surface vénusienne, de prendre des photos de cette surface. Ces photos montrent des rochers épars et découpés, preuve d'une érosion très faible.

Des faisceaux radar envoyés sur la surface de Vénus et « rebondissant » sur cette surface permettent de la « voir », de la même façon que la lumière, si on dispose d'instruments capables de détecter et d'analyser les faisceaux renvoyés, comme l'œil ou la photo le font pour la lumière. Beaucoup plus longues que les ondes lumineuses, les ondes radar « voient » moins nettement, mais c'est mieux que rien. C'est ainsi que *Pioneer Venus* a pu dresser la carte de la surface de Vénus.

La plus grande partie de cette surface est du type qui, sur Terre, correspond aux continents, plutôt qu'au fond des mers. Et tandis que la Terre possède de vastes océans (sept dixièmes de sa surface), Vénus au contraire porte un supercontinent, qui occupe les cinq sixièmes de la surface totale, le reste étant formé de petites régions analogues au fond des mers, mais évidemment sans eau.

Le supercontinent semble être assez plat, avec quelques traces de cratères, mais en petit nombre. L'atmosphère épaisse les a peut-être fait disparaître par érosion. Il y a toutefois des parties montagneuses, dont deux de très grande taille.

Dans la région du pôle Nord, Vénus possède un très grand plateau, qu'on a appelé Ishtar Terra, et qui a environ la même surface que les États-Unis. Ce plateau porte, dans sa partie orientale, une chaîne de montagnes qui a reçu le nom de monts Maxwell, dont quelques pics dépassent une altitude de 11 000 mètres au-dessus du niveau général du supercontinent. Ces pics sont nettement plus hauts que les montagnes de la Terre.

L'autre plateau est encore plus grand. Il s'étend dans la région équatoriale, et s'appelle Aphrodite Terra. Ses pics principaux sont un peu moins hauts que ceux d'Ishtar Terra.

Il est difficile de savoir si certaines des montagnes de Vénus sont des volcans. Deux d'entre elles sont peut-être des volcans, au moins des volcans éteints. L'un d'eux, le mont Rhéa, couvre une surface aussi grande que le Nouveau-Mexique.

LES SONDES VERS MERCURE

L'étude de la surface de Mercure ne pose pas autant de problèmes que celle de Vénus : il n'y a pas d'atmosphère, pas de couche de nuages. Il suffit d'envoyer une sonde dans le voisinage de la planète.

Mariner 10 a décollé le 3 novembre 1973. Le 5 février 1974, la sonde est passée tout près de Vénus, recueillant au passage des informations intéressantes, puis a continué son chemin vers Mercure.

Le 29 mars 1974, *Mariner 10* est passée à 700 kilomètres de la surface de Mercure, puis s'est placée sur orbite autour du Soleil, avec une période de 176 jours, deux fois exactement la longueur de l'« année » de Mercure. Ainsi, pendant que la sonde faisait un tour autour du Soleil, Mercure en faisait deux, et le rendez-vous a eu lieu au même point que la première fois. *Mariner 10* a frôlé ainsi deux nouvelles fois la planète, le 21 septembre 1974 et le 16 mars 1975 (cette fois à 320 kilomètres de la surface). Ce fut le dernier passage utile, car la sonde épuisa le carburant qui permettait de corriger sa trajectoire.

Lors de ses trois passages, *Mariner 10* a photographié les trois huitièmes environ de la surface de Mercure. Ces photos montrent un paysage qui ressemble beaucoup à celui de la Lune. Il y a des cratères partout, le plus grand mesurant près de 200 kilomètres de diamètre. Toutefois, Mercure a très peu de « mers ». La plus grande région presque dépourvue de cratères a environ 1 300 kilomètres de large. On l'appelle Caloris (« chaleur »), parce qu'elle est tournée vers le Soleil au point où Mercure s'approche le plus près du Soleil (*périhélie*).

La planète possède également de longues falaises, qui s'étendent sur des centaines de kilomètres, pour une hauteur de 2 500 mètres environ.

Mars

Mars est la quatrième planète à partir du Soleil, celle qui suit immédiatement la Terre. Elle est en moyenne à 230 millions de kilomètres du Soleil. Quand la Terre et Mars sont du même côté du Soleil, les deux planètes peuvent se rapprocher à 80 millions de kilomètres l'une de l'autre en moyenne. Mais l'orbite de Mars est assez elliptique, et il arrive que cette distance mesure seulement 50 millions de kilomètres. Ces rapprochements exceptionnels se produisent tous les trente-deux ans.

Le Soleil et la Lune se déplacent dans notre ciel, par rapport aux étoiles, d'une manière plus ou moins régulière de l'ouest vers l'est, mais les planètes ont des mouvements apparents beaucoup plus compliqués. La plupart du temps, d'une nuit à l'autre, elles se déplacent vers l'est par rapport aux étoiles. Mais à certains points, ce mouvement ralentit, s'arrête et repart dans l'autre sens, vers l'ouest. Ce *mouvement rétrograde* est toujours plus court que l'autre ; dans l'ensemble chaque planète se déplace vers l'est et finit par boucler un tour complet dans le ciel. De toutes les planètes, c'est Mars qui a le mouvement rétrograde le plus important.

A quoi est dû ce phénomène ? Le vieux système du monde, avec la Terre au centre, avait le plus grand mal à l'expliquer. Au contraire, le système de Copernic, avec le Soleil au centre, permet de le comprendre très

facilement. La Terre se déplace sur une orbite plus proche du Soleil que celle de Mars. Elle a moins de distance à parcourir pour faire un tour complet. Quand elle est du même côté du Soleil que Mars, elle dépasse celle-ci, de sorte que Mars *semble* aller à reculons. La comparaison du mouvement orbital de la Terre avec celui des autres planètes permet d'expliquer facilement tous les mouvements apparents rétrogrades – ce qui constitue un argument très convaincant en faveur du système héliocentrique.

Plus éloignée du Soleil que la Terre, Mars reçoit moins de soleil. C'est une petite planète, qui a seulement 6 800 kilomètres de diamètre (à peine plus que le rayon terrestre), et son atmosphère très ténue ne lui permet pas de renvoyer une fraction importante de la lumière qu'elle reçoit. Pourtant, Mars offre à l'observateur un avantage par rapport à Vénus. Quand celle-ci est à son point le plus proche de nous, elle est entre le Soleil et la Terre et nous montre sa face sombre. Mars, au contraire, quand elle est à son point le plus proche de nous, est du côté opposé au Soleil et nous montre sa face éclairée (une « pleine Mars », en somme), ce qui la rend plus brillante. Dans le meilleur des cas, Mars a une magnitude de – 2,8, ce qui en fait l'objet le plus brillant du ciel après le Soleil, la Lune et Vénus. Mais cela ne se produit que tous les trente-deux ans, lors des rapprochements exceptionnels. Et quand Mars se trouve de l'autre côté du Soleil, elle n'est pas plus brillante qu'une étoile de magnitude raisonnable.

A partir de 1580, l'astronome danois Tycho Brahé observa soigneusement Mars (sans lunette : elle n'était pas encore inventée...) pour étudier son mouvement et pouvoir prédire avec plus de précision ses mouvements à venir. Après la mort de Tycho, son assistant, l'astronome allemand Johann Kepler, utilisa ces observations pour déterminer l'orbite de Mars. Il constata qu'il devait abandonner l'idée des orbites circulaires, en vigueur parmi les astronomes depuis deux mille ans, et en 1609 il montra que les planètes devaient décrire des orbites elliptiques. Cette idée est toujours valable de nos jours, et le restera sans aucun doute définitivement.

Autre contribution décisive de Mars à l'étude du système solaire : en 1673 (nous l'avons déjà vu), Cassini détermina la parallaxe de Mars, ce qui permettait pour la première fois d'évaluer les distances entre les planètes.

Grâce à la lunette, Mars devint autre chose qu'un point lumineux. En 1659, Christiaan Huyghens y observa un dessin sombre, en forme de triangle, qu'il baptisa Syrtis Major (« grand marécage »). Grâce à ce repère, il montra que Mars tournait sur elle-même en 24,5 heures (la valeur moderne est 24,623).

Plus éloignée du Soleil que la Terre, Mars a une orbite plus longue et se déplace moins vite : il lui faut 687 jours terrestres (1,88 année terrestre) pour faire un tour, soit 668,61 jours martiens.

C'est la seule planète à avoir une période de rotation (un « jour ») aussi proche de celle de la Terre. De plus, en 1781, William Herschel montra que l'inclinaison de l'axe de Mars était très voisine de celle de l'axe de la Terre. Celui-ci est à 23,45° de la verticale, de sorte que l'hémisphère nord a son printemps et son été quand le pôle Nord penche vers le Soleil, et son automne et son hiver quand le pôle Nord penche de l'autre côté ; l'hémisphère sud, au contraire, a des saisons opposées, parce que si le pôle Nord penche vers le Soleil, le pôle Sud se tourne de l'autre côté, et réciproquement.

L'axe de Mars est incliné de 25,17° par rapport à la verticale, comme Herschel le constata en observant soigneusement la direction du déplacement des détails de sa surface à mesure que la planète tournait sur elle-même. Aussi, Mars a des saisons comme la Terre, mais chacune est presque deux fois plus longue, et, bien entendu, plus froide.

Autre ressemblance entre Mars et la Terre : en 1784, Herschel remarqua que les pôles de Mars portaient des calottes glaciaires. Dans l'ensemble, Mars ressemble plus à la Terre que n'importe quel corps céleste connu jusqu'ici. Contrairement à la Lune et à Mercure, Mars possède une atmosphère, qu'Herschel observa le premier, et ce n'est pas une atmosphère épaisse et chargée de nuages comme celle de Vénus.

Mais cette ressemblance entre Mars et la Terre ne s'étend pas à leurs satellites. La Terre en possède un gros, la Lune, contrairement à Mercure et Vénus, qui n'en ont pas. Au début, on pensait que Mars n'en avait pas non plus. En tout cas, en deux siècles et demi d'observation avec lunettes et télescopes, on n'en avait trouvé aucun.

Enfin, en 1877, lors de l'une des approches exceptionnelles de Mars, l'astronome américain Asaph Hall décida de fouiller les environs de la planète à la recherche de satellites. Puisque jusque-là on n'en avait trouvé aucun, s'ils existaient ils devaient être petits, et assez proches de la planète pour être masqués par sa lumière.

Nuit après nuit, il observa, et le 11 août 1877 il décida d'abandonner. Sa femme Angelina le poussa à essayer une nuit de plus – et c'est cette nuit-là qu'il découvrit, tout près de Mars, deux satellites minuscules. Il les baptisa Phobos et Deimos, d'après les noms des fils de Mars dans la mythologie (ces noms veulent dire « frayeur » et « terreur », et conviennent tout à fait aux fils du dieu de la guerre).

Phobos, celui des deux qui est le plus proche de la planète, n'est qu'à 9 300 kilomètres de son centre, c'est-à-dire à 6 000 kilomètres de sa surface. Il fait le tour de sa petite orbite en 7,65 heures, soit moins du tiers du temps mis par Mars pour faire un tour sur son axe. Par conséquent, Phobos va plus vite que la surface de la planète, et vu de Mars, il se lève à l'ouest et se couche à l'est. Deimos, le plus éloigné des deux satellites, est à 23 000 kilomètres du centre de Mars, et fait un tour autour de la planète en 30,3 heures.

Ces satellites étant trop petits pour apparaître, même dans les meilleurs télescopes, autrement que comme des points lumineux, il s'écoula un siècle après leur découverte sans que l'on sache autre chose à leur sujet que leur distance à la planète et leur période de révolution. Cela permit pourtant de calculer l'intensité du champ de gravitation de Mars, et par conséquent sa masse. Celle-ci vaut à peu près un dixième de la masse de la Terre, et l'accélération de la pesanteur à la surface est donc égale aux trois huitièmes de celle qui règne à la surface de la Terre : un homme pesant soixante-quinze kilos sur Terre pèserait vingt-huit kilos sur Mars.

Il s'agit pourtant d'un monde nettement plus gros que la Lune. Sa masse est 8,7 fois plus grande, et sa gravité 2,25 fois supérieure. Grosso modo, sous ce rapport, Mars est à mi-chemin entre la Terre et la Lune. Dans le cas de Vénus et Mercure, qui n'ont pas de satellites, la masse n'est pas aussi facile à évaluer. On sait maintenant que celle de Vénus vaut les quatre cinquièmes de celle de la Terre, et celle de Mercure un dix huitième seulement : à peu près deux fois moins massive que Mars, Mercure est la plus petite des huit planètes principales.

Connaissant la taille et la masse d'un objet, il est facile de calculer sa densité. Mercure, Vénus et la Terre ont des densités voisines, un peu plus de cinq fois supérieures à celle de l'eau (respectivement 5,48 5,25 et 5,52). Ces valeurs sont trop élevées pour que ces planètes puissent être constituées entièrement de roches, et on pense donc que chacune d'elles a un noyau métallique (nous en parlerons plus en détail dans le prochain chapitre).

La Lune, de densité 3,34, peut être entièrement rocheuse. Mars, intermédiaire une fois encore, doit avoir un petit noyau métallique : sa densité vaut 3,93.

LA CARTE DE MARS

Il était naturel que les astronomes tentent de dresser la carte de Mars, en représentant les taches claires et foncées visibles sur sa surface. Mais cette entreprise, facile dans le cas de la Lune, l'était beaucoup moins pour Mars, cent cinquante fois plus éloignée dans le meilleur des cas, et pourvue d'une atmosphère ténue, certes, mais capable cependant de voiler les détails de la surface.

En 1830, un astronome allemand, Wilhelm Beer, qui avait dressé une carte détaillée de la Lune, se tourna vers Mars. Il publia la première carte de la planète, montrant des zones claires et des zones foncées, où il voyait respectivement des continents et des mers. Malheureusement, d'autres astronomes tentèrent leur chance, et chacun publia une carte différente.

Parmi eux, celui qui connut le plus grand succès fut Schiaparelli (qui devait plus tard donner à la période de rotation de Mercure la valeur – fausse – de 88 jours). En 1877, lors du rapprochement exceptionnel qui permit à Hall de découvrir les deux satellites, Schiaparelli traça une carte de Mars qui différait complètement de toutes celles que l'on avait publiées auparavant. Cette fois, pourtant, tous les astronomes étaient d'accord : les télescopes s'étaient perfectionnés avec le temps, tout le monde voyait plus ou moins la même chose que Schiaparelli, et sa nouvelle carte allait durer presque un siècle. Elle donna à différentes régions des noms puisés dans la mythologie et la géographie de la Grèce antique, de Rome et de l'Égypte.

En observant Mars, Schiaparelli remarqua de fines lignes sombres, qui reliaient entre elles les taches sombres comme des détroits ou des chenaux relient nos mers. Il les appela *chenaux*, utilisant le mot italien *canali*. Malheureusement, le mot veut aussi dire *canaux*, et c'est ainsi qu'il allait être traduit dans toutes les langues. Or, les chenaux sont des structures naturelles, tandis que les canaux sont des constructions artificielles.

Les observations de Schiaparelli suscitèrent aussitôt un intérêt considérable pour la planète Mars. Bien que plus petite que la Terre et de gravité plus faible, on y voyait depuis longtemps une planète « dans le genre de la Terre ». Mais peut-être n'avait-elle pas pu garder aussi bien son atmosphère et son eau et se desséchait-elle depuis de nombreux millions d'années ? Si une vie intelligente s'y était développée, elle se battait forcément contre le dessèchement.

Aussi les gens se mirent-ils à imaginer qu'il existait une vie intelligente sur Mars et qu'elle y était plus avancée que la nôtre du point de vue technologique : les Martiens, suivant ces idées, auraient été capables de construire de gigantesques canaux pour amener l'eau des calottes glaciaires

jusqu'aux régions plus tempérées de la zone équatoriale, afin d'irriguer leurs champs.

D'autres astronomes observèrent les canaux à leur tour. Le plus enthousiaste d'entre eux fut l'Américain Percival Lowell, riche amateur qui se construisit en 1894 un laboratoire privé dans l'Arizona. Là-bas, dans l'air pur du désert en altitude, loin des villes et de leurs lumières, la visibilité était excellente, et Lowell se mit à dessiner des cartes beaucoup plus détaillées que celles de Schiaparelli. Il traça finalement plus de cinq cents canaux, et publia des livres qui popularisaient l'idée d'une vie sur Mars.

En 1897, l'écrivain anglais de science-fiction, H.G. Wells, publia en feuilleton dans un magazine le roman *La Guerre des mondes,* ce qui renforça encore la croyance générale en la vie martienne. Et quand, le 30 octobre 1938, Orson Welles réalisa une version radiophonique de la *Guerre des mondes,* avec atterrissage des Martiens dans le New Jersey, des milliers de personnes s'enfuirent épouvantées, persuadées qu'il s'agissait d'un reportage en direct...

Pourtant, beaucoup d'astronomes doutaient de la réalité des canaux de Lowell. Ne voyant pas les canaux eux-mêmes, ils s'interrogeaient sur leur existence. Maunder (qui a découvert les périodes sans taches solaires, ou « minima de Maunder ») pensa qu'il s'agissait là d'illusions d'optique. En 1913, il dessina des cercles tachetés irrégulièrement de noir et plaça des écoliers à des distances où ils distinguaient à peine ces dessins. Il leur demanda de représenter ce qu'ils voyaient, et obtint des réseaux de lignes droites qui ressemblaient de façon frappante aux canaux de Lowell...

De plus, l'observation directe semblait mettre en doute la ressemblance entre la Terre et Mars. En 1926, deux astronomes américains, William Weber Coblentz et Carl Otto Lampland, réussirent à mesurer la température de surface de Mars. Celle-ci se révéla beaucoup plus basse que prévu. En plein jour, à l'équateur et au moment où la planète était la plus proche du Soleil, la température était assez douce, mais la nuit, Mars était partout plus froide que l'Antarctique au cœur de l'hiver... La différence entre les températures du jour et de la nuit laissait supposer que l'atmosphère martienne était beaucoup plus ténue qu'on le croyait.

En 1947, l'astronome néerlando-américain Gerard Pieter Kuiper analysa le rayonnement infrarouge qui nous arrivait de Mars, et en conclut que l'atmosphère martienne était surtout formée de gaz carbonique. Il n'y trouva aucune trace d'azote, d'oxygène ou de vapeur d'eau. Il semblait donc peu probable d'y trouver des formes de vie complexes, ressemblant un tant soit peu à celles qui existaient sur la Terre. Pourtant, bien des gens se cramponnaient à leur espoir de trouver sur Mars de la végétation, et même des canaux...

LES SONDES VERS MARS

Quand les fusées commencèrent à monter au-delà de l'atmosphère, l'espoir de résoudre enfin ce problème séculaire s'éleva avec elles.

La première mission vers Mars fut réussie, avec *Mariner 4,* qui décolla le 28 novembre 1964. Le 14 juillet 1965, cette sonde passait à moins de 10 000 kilomètres de la surface de Mars. Elle prit une série de vingt photos et les traduisit en signaux radio qu'elle expédia vers la Terre. Là, ces signaux

furent traduits à nouveau, dans l'autre sens, et donnèrent des images qui montraient des cratères – mais pas de canaux.

Au moment où *Mariner 4* passait derrière Mars, ses signaux radio, avant de s'éteindre, traversaient l'atmosphère martienne, ce qui permit d'en évaluer la densité : elle était beaucoup plus faible que prévu, à peine un centième de celle de l'atmosphère terrestre.

Mariner 6 et *Mariner 7*, plus perfectionnées, sont parties respectivement le 24 février et le 27 mars 1969. Elles sont passées à 3 000 kilomètres de la surface de Mars, et ont envoyé en tout deux cents photos, couvrant une fraction appréciable de la surface totale de la planète. On y constatait que certaines régions étaient criblées de cratères, comme la Lune, tandis que d'autres étaient relativement lisses, et d'autres encore tout à fait chaotiques. Apparemment, l'évolution géologique de Mars est assez compliquée.

Mais nulle part on ne trouva la moindre trace de canaux ; l'atmosphère contient au moins 95 % de gaz carbonique, et la température est encore plus basse que ne le laissaient supposer les mesures de Coblentz et Lampland. Tout espoir semble donc vain de trouver sur Mars une vie intelligente – ou même quelque vie complexe que ce soit.

Il reste pourtant beaucoup à faire. La sonde suivante, *Mariner 9*, a atteint le voisinage de Mars le 13 novembre 1971, et au lieu de continuer tout droit, elle s'est placée sur orbite autour de la planète. C'était une chance, car peu de temps après son départ, une sorte de vent de sable s'est levé sur toute la planète, et pendant des mois les photos n'auraient pu montrer qu'une brume indistincte. Mais en orbite, la sonde a attendu la fin de la tempête, et en décembre, l'atmosphère s'est éclaircie et le travail a pu commencer.

La sonde a dressé de Mars une carte aussi détaillée que celles de la Lune, et ainsi, au bout d'un siècle, le mystère des canaux est définitivement résolu : il n'y a pas de canaux. Il s'agissait bien, comme le pensait Maunder, d'illusions d'optique. Toute la surface est sèche, et les régions sombres sont tout simplement des traînées de particules plus foncées, comme l'avait suggéré deux ans plus tôt l'astronome américain Carl Sagan.

La moitié de la planète, surtout dans l'hémisphère sud, est criblée de cratères, comme la Lune. Dans l'autre moitié, les cratères semblent avoir été estompés par le volcanisme : on y trouve de grandes montagnes qui sont sans doute des volcans, peut-être éteints depuis longtemps. Le plus grand, qui a reçu en 1973 le nom de mont Olympe, s'élève à 24 000 mètres au-dessus du niveau moyen du sol, et son cratère a un diamètre de 60 kilomètres. Il est évidemment beaucoup plus grand que n'importe quel volcan terrestre.

Il y a, à la surface de Mars, une fissure que l'on aurait bien pu prendre pour un canal. C'est une gorge immense, à laquelle on a donné le nom de Valles Marineris, et qui a environ 3 000 kilomètres de long, 500 de large, et deux de profondeur. Elle est neuf fois plus longue que le Grand Canyon du Colorado, quatorze fois plus large et deux fois plus profonde. Elle a peut-être été formée par des phénomènes volcaniques il y a deux cents millions d'années.

Il y a aussi sur Mars des traces qui serpentent, et qui ont des « affluents » qui ressemblent de façon frappante à des lits desséchés de rivières. Est-il donc possible que Mars traverse actuellement une glaciation, toute l'eau étant piégée dans les calottes polaires et dans le sous-sol ? Est-il concevable

que dans un passé relativement proche – et peut-être dans un avenir pas trop lointain – des conditions moins défavorables aient permis, ou permettent à nouveau, la présence d'eau liquide et de rivières ? Si oui, ne pourrait-on pas trouver dans le sol martien des formes de vie primitive, résistant tant bien que mal aux conditions actuelles ?

De toute évidence, un atterrissage en douceur s'imposait. *Viking 1* et *Viking 2* ce sont envolées respectivement le 20 août et le 9 septembre 1975. *Viking 1* s'est placée en orbite autour de Mars le 19 juin 1976 et a envoyé vers le sol un module d'atterrissage, qui s'est posé en douceur le 20 juillet. Quelques semaines plus tard, le module de *Viking 2* a atterri un peu plus au nord.

Pendant leur passage à travers l'atmosphère martienne, les modules d'atterrissage l'ont analysée et y ont trouvé, en plus du gaz carbonique, 2,7 % d'azote, 1,6 % d'argon et des traces d'oxygène.

Au sol, ils ont mesuré une température diurne maximale de l'ordre de -30 °C. Il ne semble y avoir aucune chance que la température atteigne jamais, où que ce soit sur Mars, le point de fusion de l'eau, ce qui exclut la présence d'eau liquide. Il fait trop froid pour la vie, de même que sur Vénus il faisait trop chaud. Tout au moins, sur Mars, la température est trop faible pour toutes les formes de vie, sauf peut-être les plus simples. Le froid est tel que le gaz carbonique lui-même est solide dans les régions les plus froides et semble former au moins une partie des calottes glaciaires des pôles.

Les modules d'atterrissage ont envoyé des photos de la surface de Mars, et analysé son sol. Il se compose à 80 % d'argile riche en fer, et celui-ci est peut-être sous forme de *limonite*, le dérivé du fer qui donne aux briques leur couleur rouge. La couleur rougeâtre de Mars, qui effrayait les Anciens parce qu'ils l'associaient au sang, n'a rien à voir avec celui-ci : Mars est tout simplement une planète rouillée...

Les laboratoires chimiques avec lesquels les modules ont analysé le sol ont été spécialement équipés pour déceler la présence de cellules vivantes. Or, les trois expériences réalisées n'ont pas donné de résultat net : la vie existait peut-être, mais on n'en était pas sûr. Ce qui causa cette incertitude, c'était l'absence de quantités détectables des composés organiques, ceux que nous avons l'habitude de considérer comme les matériaux de la vie. Les scientifiques se sont refusés à envisager une vie fondée sur d'autres matériaux, et la solution de ce problème devra donc attendre l'atterrissage de modules plus perfectionnés, ou mieux, d'équipes humaines.

LES SATELLITES DE MARS

Au départ, on n'avait pas prévu d'étude détaillée des satellites par les sondes martiennes. Mais pendant la tempête de sable, *Mariner 9* n'a rien eu de mieux à faire que de filmer les satellites. Ceux-ci se sont révélés de forme irrégulière. (On imagine généralement comme des sphères tous les corps astronomiques, mais cette forme n'est obligatoire que dans le cas de ceux qui sont assez massifs pour que leur gravitation leur impose une forme régulière.) En fait, chacun de ces satellites ressemblait à une « pomme de terre au four », et certains cratères ressemblaient d'une façon étonnante aux « yeux » de la pomme de terre.

Les dimensions de Phobos, le plus grand des deux, sont de vingt à vingt-sept kilomètres, et celles de Deimos, de dix à seize kilomètres. Ce sont des montagnes qui « volent » autour de Mars. Leur grand axe pointe toujours vers Mars, par un « verrouillage » analogue à celui de la Terre et de la Lune.

On a donné aux deux plus grands cratères de Phobos les noms de Hall et Stickney, en l'honneur de celui qui a découvert les satellites de Mars et en l'honneur de sa femme, qui l'a poussé à persévérer une nuit de plus. Les deux plus grands cratères de Deimos s'appellent Voltaire et Swift, du nom des deux écrivains, qui avaient tous deux imaginé que Mars possédait deux satellites...

Jupiter

Cinquième planète à partir du Soleil, Jupiter est le géant du système solaire. Son diamètre est de 140 000 kilomètres, soit 11,2 fois celui de la Terre. Sa masse vaut 318,4 fois celle de la Terre. La planète géante, à elle seule, pèse deux fois plus que toutes les autres ensemble. Pourtant, elle est toute petite par rapport au Soleil, dont la masse est encore mille fois plus grande.

En moyenne, Jupiter se trouve à 780 millions de kilomètres du Soleil (5,2 fois plus loin que la Terre). La distance entre les deux planètes ne descend jamais en dessous de 620 millions de kilomètres, même quand elles sont du même côté du Soleil. Quant à la lumière reçue du Soleil, elle est vingt-sept fois moins intense sur Jupiter que sur la Terre. Pourtant, la planète est si grande qu'elle brille vivement dans notre ciel.

Sa magnitude maximale est de -2,5, ce qui correspond à un éclat beaucoup plus vif que celui de n'importe quelle étoile. Vénus et Mars, à leur maximum, dépassent Jupiter en éclat (Vénus la dépasse même considérablement). Mais Vénus et Mars sont la plupart du temps beaucoup plus pâles, quand elles sont de l'autre côté de leur orbite. Jupiter, au contraire, pâlit à peine quand elle s'éloigne de la Terre : son orbite est si grande qu'il importe peu qu'elle soit ou non du même côté que nous du Soleil. Jupiter est donc souvent l'objet le plus brillant du ciel (après le Soleil et la Lune), d'autant plus que – contrairement à Vénus – il lui arrive de rester visible toute la nuit. C'est donc à juste titre qu'on lui a donné le nom de roi des dieux dans la mythologie gréco-romaine.

LES SATELLITES DE JUPITER

Quand Galilée tourna sa première lunette vers le ciel, il ne tarda pas à la pointer vers Jupiter. Le 7 janvier 1610, il remarqua près de la planète trois points lumineux bien alignés avec elle, deux d'un côté et un de l'autre. Les nuits suivantes, ils étaient toujours là, mais dans des positions différentes, tantôt à droite, tantôt à gauche de Jupiter. Le 13 janvier, Galilée en découvrit un quatrième.

Il se rendit compte que quatre corps célestes tournaient autour de Jupiter, comme la Lune autour de la Terre. C'étaient les premiers objets

du système solaire, invisibles à l'œil nu, que la lunette permettait de découvrir. D'autre part, ils étaient la preuve manifeste que tous les objets célestes ne sont pas en orbite autour de la Terre.

Pour les désigner, Kepler inventa le mot *satellite*, à partir d'un mot latin désignant ceux qui s'agitent dans l'entourage d'un personnage important. Depuis lors, on a donné ce nom à tous les objets qui tournent autour d'une planète, qu'ils soient naturels (comme la Lune) ou artificiels (comme *Spoutnik I*).

Ces quatre satellites de Jupiter sont regroupés sous le nom de *satellites galiléens*, mais peu après leur découverte par Galilée, ils ont été baptisés individuellement par l'astronome hollandais Simon Marius. A partir de Jupiter, on a trouvé ainsi successivement Io, Europe, Ganymède et Callisto (ce sont les noms de quatre personnes souvent associées à Jupiter dans la mythologie).

Le plus proche de Jupiter, Io, est à 420 000 kilomètres du centre de la planète, distance analogue à celle qui sépare la Lune du centre de la Terre. Mais tandis que la Lune fait un tour en 27,3 jours, Io tourne autour de Jupiter en 1,77 jour. En effet, le champ de gravitation de Jupiter est beaucoup plus intense que celui de la Terre, à cause de son énorme masse. On peut d'ailleurs calculer cette masse précisément à partir de la période de révolution d'Io.

Europe, Ganymède et Callisto sont à 670 000, 1 060 000 et 1 890 000 kilomètres de Jupiter, et en font respectivement le tour en 3,55 jours, 7,16 jours et 16,7 jours. Ainsi, Jupiter et ses quatre satellites forment un système solaire en miniature, dont la découverte devait faire tomber bien des préjugés contre le système de Copernic.

Ces résultats ont permis de calculer la masse de Jupiter, et on lui a trouvé une valeur étonnamment faible. Bien sûr, cette masse vaut 318,4 fois celle de la Terre, mais le volume de Jupiter vaut 1 400 fois celui de la Terre. Pourquoi donc sa masse n'est-elle pas 1 400 fois la masse terrestre ? C'est forcément parce que la densité de Jupiter est plus faible : elle vaut seulement 1,34, c'est-à-dire le quart de celle de la Terre. Jupiter est donc constituée de matériaux moins denses que les roches et les métaux.

Quant aux satellites, ils sont comparables à notre Lune. Europe, le plus petit des quatre, a un diamètre de 3 100 kilomètres, un peu moins que la Lune. Io, avec un diamètre de 3 650 kilomètres, a presque les dimensions de la Lune. Callisto et Ganymède sont plus grands, avec des diamètres respectifs de 4 850 et 5 250 kilomètres.

Ganymède est le plus gros satellite du système solaire, et sa masse vaut 2,5 fois celle de la Lune. En fait, Ganymède est nettement plus gros que Mercure, dont la taille est à peu près celle de Callisto. Mais Mercure est constituée de matériaux plus denses, et Ganymède, bien plus gros, n'a que les trois cinquièmes de sa masse. Les deux satellites intérieurs, Io et Europe, ont à peu près la densité de la Lune, et sont sans doute formés de roches. Ganymède et Callisto ont des densités de l'ordre de celle de Jupiter et sont formés de matériaux plus légers.

Il n'est pas étonnant que Jupiter ait quatre gros satellites et la Terre un seulement, compte tenu de la taille bien supérieure de la planète géante. En fait, ce qui est étonnant, c'est que Jupiter n'en ait pas encore plus, ou la Terre encore moins.

Les quatre satellites galiléens ont ensemble une masse qui vaut 6,2 fois celle de la Lune, mais seulement 1/4 200 fois celle de Jupiter, leur planète.

La Lune, à elle toute seule, possède 1/81 fois la masse de sa planète, la Terre.

Le plus souvent, les satellites sont minuscules en comparaison des planètes autour desquelles ils tournent. Mercure et Vénus n'en ont pas du tout (bien que Vénus soit presque aussi grosse que la Terre) et Mars a deux satellites, mais tout à fait insignifiants. Celui de la Terre est si gros que les deux forment ce que l'on pourrait presque appeler une « planète double » (jusqu'à une date très récente, on croyait que c'était le seul exemple de ce genre, mais à tort, comme nous le verrons bientôt dans le même chapitre).

Pendant près de trois siècles après la découverte de Galilée, on ne découvrit aucun nouveau satellite de Jupiter. Pourtant, pendant cette période, on en découvrit quinze autour d'autres planètes.

Enfin, en 1892, l'astronome américain Edward Emerson Barnard détecta un petit point lumineux près de Jupiter, si faible qu'on le distinguait à peine si près de la planète brillante. C'était un cinquième satellite de Jupiter, et le dernier satellite découvert par observation visuelle. Depuis lors, on en a découvert d'autres, mais par photographie, soit à partir de la Terre, soit à partir des sondes spatiales.

Ce cinquième satellite a reçu le nom d'Amalthée (la chèvre qui aurait allaité Jupiter enfant), mais ce nom n'est devenu officiel que dans les années soixante-dix.

Amalthée n'est qu'à 180 000 kilomètres du centre de Jupiter et en fait le tour en 11,95 heures. Ce satellite est plus proche que les quatre galiléens, ce qui explique en partie sa découverte tardive : il se perd dans l'éclat de Jupiter. D'autre part il est très petit (250 kilomètres de diamètre) et ne renvoie pas beaucoup de lumière.

Pourtant, on allait s'apercevoir que Jupiter avait beaucoup d'autres satellites, encore plus petits qu'Amalthée, et par conséquent encore moins brillants. La plupart se trouvent loin de la planète, bien au-delà de l'orbite de Callisto, le plus extérieur des satellites galiléens. On en a détecté huit entre 1904 et 1974. On a commencé par se contenter de leur donner à chacun un numéro, en chiffres romains, indiquant l'ordre de leur découverte, de Jupiter VI à Jupiter XIII.

L'astronome américain Charles Dillon Perrine a découvert Jupiter VI en décembre 1904 et Jupiter VII en janvier 1905. Le premier a un diamètre de 100 kilomètres, le second de 30 kilomètres seulement.

C'est un astronome anglais, P. J. Melotte, qui a découvert Jupiter VIII en 1908. Les quatre suivants s'inscrivent à l'actif de l'Américain Seth B. Nicholson : Jupiter IX en 1914, X et XI en 1938, et XII en 1951. Tous mesurent environ 20 kilomètres.

Enfin, le 10 septembre 1974, l'astronome américain Charles T. Kowal a découvert Jupiter XIII qui ne mesure que 16 kilomètres.

Ces satellites extérieurs se divisent en deux groupes. Quatre d'entre eux (VI, VII, X et XIII) sont à peu près à 11 millions de kilomètres de Jupiter, c'est-à-dire environ six fois plus loin que Callisto, le plus extérieur des quatre galiléens. Les quatre derniers sont deux fois plus loin encore, environ à 22 millions de kilomètres de la planète géante.

Les satellites galiléens tournent tous dans le plan équatorial de Jupiter, et sur des orbites pratiquement circulaires. C'est un résultat prévisible de l'attraction de Jupiter, comme nous le verrons plus en détail au chapitre suivant, à propos des marées. Si un satellite tourne dans un autre plan

(on dit qu'il est *incliné*), ou si son orbite est excentrée, l'effet de marée, peu à peu, attire ce satellite dans le plan équatorial et rend son orbite circulaire.

Or, cet effet de marée, s'il est proportionnel à la masse de la planète, décroît très vite en fonction de la distance, et il est proportionnel à la taille de l'objet affecté. Par conséquent, en dépit de sa masse énorme, Jupiter n'exerce qu'un effet de marée très faible sur les petits satellites extérieurs. Aussi, bien que les deux groupes de quatre aient chacun des distances très voisines au centre de la planète, il n'y a pour l'instant aucun danger de collision, car leurs orbites sont différemment inclinées et différemment excentrées.

Les satellites du groupe le plus extérieur ont des orbites si inclinées qu'elles sont pour ainsi dire « retournées » : ils tournent autour de la planète d'une façon rétrograde, dans le sens des aiguilles d'une montre (vus depuis un point situé au-dessus du pôle Nord de Jupiter), au lieu de tourner dans l'autre sens, comme tous les autres satellites de Jupiter.

Ces petits satellites extérieurs sont peut-être des astéroïdes (nous en parlerons un peu plus loin) capturés par l'attraction de Jupiter, ce qui expliquerait l'irrégularité de leurs orbites, l'effet de marée n'ayant pas eu le temps, depuis leur capture, de les « mettre au pas ». De plus, on peut montrer qu'une planète capture plus facilement un objet si celui-ci se présente de manière à s'engager sur une orbite rétrograde.

Le satellite qui s'éloigne le plus de Jupiter est Jupiter VIII, baptisé maintenant Pasiphae (tous les satellites extérieurs ont désormais, en plus de leur numéro, un nom tiré de la mythologie). Son orbite est si excentrée qu'à sa position la plus éloignée, Pasiphae passe à 33 millions de kilomètres de Jupiter, soit 80 fois la distance Terre-Lune. C'est jusqu'ici un record dans le système solaire.

Jupiter IX (Sinope) tourne en moyenne un peu plus loin que Pasiphae, et met donc un peu plus longtemps pour faire un tour. Sa période de révolution est 758 jours (plus de deux ans), ce qui est également un record du système solaire.

LA FORME ET LA SURFACE DE JUPITER

Et la planète elle-même ? En 1691, Cassini observa Jupiter à la lunette, et remarqua que ce n'était pas un cercle parfait, mais plutôt une ellipse. Cela voulait dire que la planète, au lieu d'être sphérique, avait une forme aplatie, comme celle d'une mandarine.

Ce résultat paraissait étonnant, car le Soleil et la Lune (au moins quand elle est pleine...) apparaissaient comme des cercles parfaits et étaient donc sans doute des sphères parfaites. Toutefois, la théorie de Newton, alors fort récente, expliquait parfaitement le phénomène. Comme nous le verrons dans le prochain chapitre, une sphère *en rotation* doit être aplatie. Sa rotation provoque un renflement à l'équateur et un aplatissement aux pôles, d'autant plus nets que la rotation est plus rapide.

Aussi, le *diamètre équatorial* (c'est-à-dire mesuré entre deux points diamétralement opposés sur l'équateur) doit-il être plus grand que le *diamètre polaire* (la distance entre les pôles). La différence des deux, divisée par le diamètre équatorial, donne un nombre qui caractérise l'aplatissement de la planète. Dans le cas de Jupiter, les diamètres valent

respectivement 143 000 kilomètres (diamètre équatorial) et 134 000 kilomètres (diamètre polaire). La différence, 9 000 kilomètres, vaut plus des deux tiers du diamètre de la Terre, et si on la divise par le diamètre équatorial, on trouve un aplatissement de 0,062, soit environ un seizième.

Mercure, Vénus et notre Lune, qui tournent lentement sur elles-mêmes, n'ont pas d'aplatissement perceptible. Le Soleil tourne relativement vite, mais son énorme champ gravitationnel l'empêche de s'aplatir beaucoup, et lui aussi a un aplatissement trop petit pour être mesurable. La Terre tourne plus vite, et son aplatissement vaut 0,0033. Mars, qui tourne assez vite aussi, et qui a un champ de gravitation plus faible, a un aplatissement un peu supérieur : 0,0052.

L'aplatissement de Jupiter est près de vingt fois supérieur à celui de la Terre, bien que son champ de gravitation soit considérablement plus fort. On peut donc s'attendre à ce que sa vitesse de rotation soit très grande. Effectivement, Cassini mesura cette vitesse, en observant le déplacement des détails de la surface : la planète tournait sur elle-même en un peu moins de dix heures (la valeur moderne est 9,85 heures, soit les deux cinquièmes d'un jour terrestre).

Or, non seulement Jupiter tourne plus vite que la Terre, mais c'est une planète beaucoup plus grosse. Aussi, un point de l'équateur terrestre se déplace à 1 700 km/h, et un point de l'équateur de Jupiter à 45 000 km/h !

Les taches remarquées à la surface par Cassini (et par d'autres après lui) changeaient constamment, et ne faisaient donc pas partie d'une surface solide. Il s'agissait probablement, comme dans le cas de Vénus, d'une couche de nuages, et les taches étaient sans doute des ouragans. La surface présente aussi des lignes parallèles à l'équateur, qui révèlent sans doute des vents dominants. Dans l'ensemble, Jupiter est jaunâtre, les lignes allant de l'orange au brun, avec de temps en temps un peu de blanc, de bleu ou de gris.

Le détail le plus remarquable de la surface de Jupiter fut signalé pour la première fois par l'Anglais Robert Hooke en 1664, et en 1672, Cassini le représenta comme une grande tache ronde sur un dessin de Jupiter. On le retrouva sur d'autres dessins, les années suivantes, mais c'est seulement en 1878 qu'il fut décrit en détail par un astronome allemand, Ernst Wilhelm Tempel. Comme sa couleur lui semblait très rouge, il lui donna le nom, qui lui restera, de la Grande Tache rouge. Sa couleur est variable, et parfois si pâle qu'il faut un bon télescope pour la distinguer. C'est une ellipse qui mesure 45 000 kilomètres d'est en ouest et 13 000 du nord au sud.

Les astronomes s'interrogèrent sur sa nature. Pour certains, c'était un immense ouragan. Pour d'autres, Jupiter était si massive que sa température était beaucoup plus élevée que celle des autres planètes – et peut-être suffisante pour atteindre celle du fer rouge... Mais si Jupiter est sans doute très chaude à l'intérieur, sa surface ne l'est pas : en 1926, un astronome américain, Donald Howard Menzel, a montré que la température de la couche de nuages que nous voyons était de -135 °C.

LES MATÉRIAUX DE JUPITER

Pour avoir une densité aussi faible, Jupiter doit être riche en matériaux plus légers que les métaux ou les roches.

Or, les matériaux les plus communs dans l'Univers en général sont l'hydrogène et l'hélium. 90 % des atomes sont des atomes d'hydrogène, et 9 % des atomes d'hélium. Ce n'est pas étonnant : l'atome d'hydrogène est le plus simple de tous, et le suivant est l'atome d'hélium. Le reste (1 % des atomes de l'Univers) est surtout formé de carbone, d'oxygène, d'azote, de néon et de soufre. Les atomes d'hydrogène et d'oxygène se combinent pour former des molécules d'eau, hydrogène et carbone forment du méthane, hydrogène et azote de l'ammoniac.

La densité de tous ces corps, dans des conditions normales, est inférieure ou égale à celle de l'eau. Mais à des pressions élevées, comme celles qui doivent régner à l'intérieur de Jupiter, ces densités augmentent, et peuvent devenir supérieures à celle de l'eau. Si Jupiter était formé de ces matériaux, cela rendrait bien compte de sa faible densité.

En 1932 un astronome allemand, Rupert Wildt, a étudié la lumière renvoyée par Jupiter, et constaté que certaines longueurs d'onde y étaient absentes – celles qui sont absorbées par l'ammoniac et le méthane. Il en conclut que ces deux substances, au moins, formaient une partie de l'atmosphère de Jupiter.

En 1952 Jupiter devait passer devant l'étoile Sigma du Bélier, événement que deux astronomes américains, William Alvin Baum et Arthur Dode Code, s'apprêtaient à étudier avec soin. Quand l'étoile s'approcha de Jupiter, sa lumière traversa l'atmosphère ténue qui surmonte la couche de nuages. L'étude de la lumière transmise a montré que cette partie de l'atmosphère se composait surtout d'hydrogène et d'hélium. En 1963, un autre astronome américain, Hyron Spinrad, y a trouvé également du néon.

Tous ces corps sont gazeux dans les conditions terrestres, et s'ils constituent une part importante de la masse de Jupiter, il semble juste de considérer cette planète comme une *géante gazeuse*

Les premières sondes lancées vers Jupiter, *Pioneer 10* et *Pioneer 11*, ont décollé respectivement le 2 mars 1972 et le 5 avril 1973. *Pioneer 10* est passée le 3 décembre 1973 à 130 000 kilomètres de la surface visible de Jupiter. Un an plus tard, *Pioneer 11* a survolé, à 40 000 kilomètres seulement, le pôle Nord de la planète, que les hommes ont vu ainsi pour la première fois.

Les deux sondes suivantes, plus perfectionnées, étaient *Voyager 1* et *Voyager 2*. Lancées respectivement le 20 août et le 5 septembre 1977, elles ont frôlé Jupiter en mars et en juillet 1979.

Ces sondes ont confirmé les premiers résultats sur l'atmosphère de Jupiter. Elle est surtout formée d'hydrogène et d'hélium, dans une proportion de dix pour un (à peu près la proportion générale de l'Univers). Les constituants nouveaux, pas encore détectés depuis la Terre, comprennent l'acétylène et l'éthylène (tous deux formés de carbone et d'hydrogène), l'eau, l'oxyde de carbone, et des composés du phosphore et du germanium.

Il est évident que l'atmosphère de Jupiter est le siège de phénomènes chimiques complexes, et que nous manquerons d'informations à son sujet jusqu'à ce qu'une sonde y pénètre – et y survive assez longtemps pour pouvoir nous envoyer des mesures. La Grande Tache rouge, elle, est bien – comme le supposaient la plupart des astronomes – un ouragan gigantesque (beaucoup plus vaste que la Terre entière) et pratiquement permanent.

La planète semble entièrement liquide. La température s'élève rapidement vers l'intérieur, et la pression contribue à faire de l'hydrogène un

liquide chauffé « au rouge ». Au centre, il y a peut-être un noyau extrêmement chaud d'hydrogène solide, sous la forme appelée « hydrogène métallique ». Mais les conditions qui règnent dans l'intérieur de Jupiter sont difficiles à reproduire en laboratoire, et il s'écoulera peut-être beaucoup de temps avant que nous puissions avoir une idée précise de ce qui s'y passe.

LES SONDES VERS JUPITER

Leur premier travail a été de photographier en gros plan les quatre satellites galiléens. Pour la première fois, les hommes ont pu les voir autrement que comme de minuscules disques lumineux.

On a pu ainsi préciser les idées au sujet de leur taille et de leur masse, qui se sont révélées assez proches des estimations calculées depuis la Terre. Toutefois, Io (le plus intérieur des quatre satellites galiléens) est plus massif (de 25 %) qu'on ne le pensait.

Ganymède et Callisto, comme leur faible densité le laissait prévoir, sont formés de substances légères, comme l'eau. Aux températures faibles qu'on attend sur des objets aussi éloignés du Soleil – et trop petits pour posséder la chaleur interne de Jupiter, ou même de la Terre – ces substances sont sous forme solide, et on les désigne généralement sous le nom de « glace ». Les deux satellites sont criblés de cratères.

Un facteur qui pourrait échauffer les satellites de Jupiter est le frottement engendré par les déformations provoquées par les marées dues au voisinage de la planète géante. Mais Ganymède et Callisto sont suffisamment éloignés pour que cet effet soit insignifiant, et ils restent glacés.

Europe est plus près de Jupiter, et a dû traverser une phase assez chaude pour l'empêcher d'accumuler beaucoup de glace, ou pour faire fondre et vaporiser celle qu'il possédait, et donc la disperser dans l'espace, car la masse d'Europe est trop faible pour lui permettre de garder une atmosphère. C'est peut-être ce manque de glace qui rend Europe et Io beaucoup plus petits que Ganymède et Callisto.

Toutefois, Europe a gardé assez de glace pour posséder un océan qui couvre toute sa surface (comme on pensait autrefois en découvrir un sur Vénus). A la température d'Europe, cet océan est gelé, et il s'agit d'un glacier à l'échelle d'un monde. Qui plus est, ce glacier est remarquablement lisse (Europe est le corps le plus « poli » découvert à ce jour), bien que marqué par un réseau de lignes sombres qui évoquent d'une façon frappante les vieilles cartes de Mars dressées par Lowell.

Si ce glacier est si lisse et dépourvu de cratères, c'est peut-être parce qu'il recouvre de l'eau liquide, fondue par échauffement dû à l'effet de marée. Ainsi, une météorite suffisamment grosse casserait la couche de glace, mais l'eau viendrait geler à nouveau et boucher le trou. Des chocs moins violents causeraient des fissures passagères, de même d'ailleurs que l'effet de marée, ou d'autres facteurs. Dans l'ensemble, toutefois, la surface est lisse.

Io, celui des satellites galiléens qui est le plus proche de Jupiter, est le plus soumis à l'échauffement par effet de marée et semble parfaitement sec. C'est un satellite qui posait déjà des problèmes ardus avant le voyage des sondes. En 1974, l'astronome américain Robert Brown avait déjà signalé que ce satellite était entouré d'un brouillard jaune de sodium, qui

semblait s'étendre sur toute son orbite, comme un anneau autour de Jupiter. Ce brouillard devait venir du satellite lui-même, mais personne ne savait pourquoi.

Les sondes Pioneer ont montré qu'Io possèdait une atmosphère très ténue, dont la densité était 1/20 000 celle de l'atmosphère terrestre. Mais ce sont les sondes Voyager qui ont donné la clef du mystère : Io possède des volcans actifs – les seuls trouvés jusqu'ici ailleurs que sur la Terre. Il semble que des zones de roche fondue (sans doute par les frottements dus à l'effet de marée de Jupiter) s'étendent sous la surface, et qu'à plusieurs endroits cette lave ait crevé la croûte, expulsant des jets de sodium et de soufre, qui rendent compte à la fois de l'atmosphère d'Io et de l'anneau de brouillard où voyage ce satellite. Sa surface est d'ailleurs couverte de soufre, qui lui donne des tons allant du jaune au brun. On y trouve peu de cratères, sans doute rapidement comblés par la lave. Seuls quelques-uns sont assez récents pour qu'on en distingue encore la trace.

Plus près encore de Jupiter, on rencontre Amalthée, qui n'est visible de la Terre que comme un faible point lumineux. Les sondes Voyager nous ont montré ce satellite comme une sorte d'éponge irrégulière, comme les deux satellites de Mars, mais en beaucoup plus gros : les dimensions d'Amalthée varient entre 250 et 110 kilomètres.

Les sondes ont encore découvert trois nouveaux satellites, tous plus proches de Jupiter qu'Amalthée, et considérablement plus petits. Ce sont Jupiter XIV, Jupiter XV et Jupiter XVI, de diamètres respectifs 22, 80 et 40 kilomètres. Compte tenu de leur petite taille et du fait qu'ils sont éblouis par Jupiter parce qu'ils en sont proches, ils sont tout à fait invisibles depuis la Terre.

Jupiter XVI est le plus proche de la planète, à une distance de 130 000 kilomètres au centre de la planète, c'est-à-dire 60 000 kilomètres seulement au-dessus de la surface. Il fait un tour autour de Jupiter en 7,07 heures. A peine plus éloigné, Jupiter XIV a une période de révolution de 7,13 heures. Ces deux satellites tournent autour de Jupiter plus vite que la planète sur elle-même, et si l'on pouvait les observer à partir des nuages de Jupiter, on les verrait se lever à l'ouest et se coucher à l'est, comme le fait Phobos vu de Mars.

A l'intérieur de l'orbite du satellite le plus proche de Jupiter, un certain nombre de débris forment un anneau mince et clairsemé autour de la planète. Cet anneau est trop ténu pour être visible de la Terre.

Saturne

C'est la planète la plus éloignée qu'aient connue les Anciens, car en dépit de la distance considérable qui nous en sépare, son éclat est très vif. Sa magnitude maximale vaut –0,75, c'est-à-dire que son éclat est plus vif que celui de n'importe quelle étoile, à part Sirius. Saturne est aussi plus brillante que Mercure, et de toute façon plus facile à observer, puisqu'étant plus éloignée que nous du Soleil, cette planète ne reste pas dans le voisinage apparent de notre étoile, mais peut briller dans le ciel de minuit.

Sa distance moyenne au Soleil vaut 1 420 000 000 kilomètres, près de deux fois plus que pour Jupiter. Saturne fait le tour du Soleil en 29,5

années, au lieu de 11,8 pour Jupiter. L'année saturnienne vaut donc deux fois et demie celle de Jupiter.

Sous bien des rapports, Saturne fait un peu figure de « parent pauvre » de Jupiter. Sa taille, par exemple, en fait la deuxième planète du système. Son diamètre équatorial vaut en effet 120 000 kilomètres, les cinq sixièmes environ de celui de Jupiter. C'est cette taille plus faible, et surtout un éloignement plus grand du Soleil, qui rend Saturne moins brillante que Jupiter. Mais elle fait tout de même assez bonne figure.

Sa masse vaut 95,1 fois celle de la Terre, ce qui en fait la planète la plus lourde du système – après Jupiter, bien entendu ! Sa masse ne vaut que les trois dixièmes de celle de Jupiter, alors que son volume atteint les six dixièmes de celui de la planète géante.

Pour avoir une masse relativement si faible dans un volume si grand, il faut vraiment que Saturne ait une densité très petite. En fait, c'est le corps le moins dense du système solaire : 0,7 fois la densité de l'eau ! Si on pouvait empaqueter Saturne dans du plastique, pour l'empêcher de se dissoudre ou de se disperser, et si on trouvait un océan assez vaste, en y plaçant Saturne, la planète y flotterait. Selon toute vraisemblance, Saturne est encore plus riche en hydrogène – et encore plus pauvre en quoi que ce soit d'autre – que Jupiter. De plus, son champ de gravitation est moins fort, et comprime moins les matériaux.

Saturne tourne assez vite autour de son axe, mais – malgré sa taille plus faible – moins vite que Jupiter : sa période de rotation vaut 10,67 heures terrestres, de sorte que son « jour » vaut 1,08 fois celui de Jupiter.

Bien que Saturne tourne sur son axe moins vite que Jupiter, sa forme est encore plus aplatie. En effet, ses couches externes sont moins denses, et son champ gravitationnel moins intense. Saturne est donc, de tous les corps du système solaire, celui qui possède le renflement équatorial le plus important, et l'aplatissement le plus fort : 0,102 soit 1,6 fois plus que Jupiter et 30 fois plus que la Terre. Son diamètre équatorial est de 120 000 kilomètres et son diamètre polaire de 108 000 kilomètres seulement : la différence, 12 000 kilomètres, est presque égale au diamètre de la Terre.

LES ANNEAUX DE SATURNE

Sous ce rapport, Saturne est unique – d'une beauté unique. Quand Galilée pointa pour la première fois sa lunette, bien imparfaite, vers Saturne, la forme de la planète le surprit : il crut y voir une boule centrale, flanquée de deux autres plus petites. Il poursuivit ses observations, mais les deux petites boules devenaient de plus en plus indistinctes, et finalement, vers la fin de l'année 1612, elles disparurent complètement.

D'autres astronomes constatèrent aussi quelque chose de bizarre à propos de Saturne, mais c'est seulement en 1656 que Christiaan Huygens expliqua correctement le phénomène : Saturne était entouré d'un anneau brillant qui n'était nulle part en contact avec la planète.

Comme celui de la Terre, l'axe de rotation de Saturne est incliné : son inclinaison est de 26,73 ° (23,45 pour celui de la Terre). Quand Saturne est à un bout de son orbite, nous voyons de dessus le côté le plus proche de l'anneau ; l'autre côté est caché par la planète. Quand Saturne est à l'autre bout de son orbite, nous voyons de dessous le côté le plus proche de l'anneau, l'autre côté étant caché par la planète. Saturne met un peu

plus de quatorze ans pour aller d'un bout à l'autre de son orbite. Pendant ce temps, l'anneau nous semble passer progressivement de sa position basse (où nous le voyons par dessus) à sa position haute (où nous le voyons par dessous). A mi-chemin, nous le voyons exactement par la tranche – c'est-à-dire que nous ne le voyons pas, car il est trop mince pour cela. Ainsi, tous les quatorze ans, l'anneau disparaît pendant quelque temps. C'est ce qui s'est passé à la fin de 1612, lors des observations de Galilée. Et on raconte qu'il en a été si déçu qu'il n'a plus jamais observé Saturne...

En 1675, Cassini remarqua que l'anneau n'était pas uniformément lumineux : une ligne sombre le partageait en deux anneaux concentriques, un anneau intérieur, et un anneau extérieur moins large et moins brillant que le premier. Depuis cette date, on parle au pluriel des anneaux de Saturne, et la ligne noire qui les sépare a reçu le nom de « division de Cassini ».

En 1826, l'astronome germano-russe Friedrich G.W. von Struve appela respectivement A et B l'anneau extérieur et l'anneau intérieur. En 1850, un astronome américain, William Cranch Bond, découvrit un anneau peu lumineux, encore plus près de la planète. Cet anneau C n'est séparé de l'anneau B par aucune division nette.

Il n'y a rien de comparable aux anneaux de Saturne dans le système solaire. Bien sûr, on sait maintenant qu'un anneau ténu entoure Jupiter, et peut-être toutes les géantes gazeuses comme Jupiter et Saturne ont-elles un anneau de débris tout près de leur surface. Mais celui de Jupiter est tout à fait insignifiant, alors que les anneaux de Saturne sont splendides. La partie visible depuis la Terre mesure d'un bout à l'autre 270 000 kilomètres, soit 21 fois le diamètre de la Terre, ou presque deux fois celui de Jupiter.

Quelle est la nature des anneaux de Saturne ? Cassini y voyait de minces disques rigides et lisses. Mais en 1785, Laplace (qui allait proposer plus tard l'hypothèse de la nébuleuse) fit remarquer que les différentes parties des anneaux, étant situées à des distances différentes du centre de la planète, étaient soumises à des champs de gravitation différents. Ceci pouvait faire voler l'anneau en éclats. Laplace suggéra qu'il s'agissait en fait d'un grand nombre d'anneaux minces, si rapprochés qu'ils semblaient, depuis la Terre, former une surface continue.

En 1855, toutefois, Maxwell (qui allait être plus tard mondialement connu pour sa théorie des ondes électromagnétiques) montra que même cette solution n'était pas satisfaisante. La seule façon pour les anneaux d'échapper à la destruction par effet de marée était d'être constitués de petits blocs innombrables, répartis de manière à donner – depuis la Terre – l'impression d'anneaux continus. Cette idée est maintenant universellement admise.

Abordant le problème sous un autre angle, un astronome français, Edouard Roche, montra que tout solide qui s'approche trop près d'un corps considérablement plus gros subit des forces de marée d'une puissance telle qu'il peut voler en éclats. La distance fatale, en dessous de laquelle cela se produit, s'appelle la *limite de Roche,* et vaut 2,44 fois le *rayon équatorial* de la planète (la distance entre le centre et un point de l'équateur).

Ainsi, pour Saturne, la limite de Roche vaut 2,44 fois le rayon équatorial (60 000 kilomètres), soit environ 140 000 kilomètres. Or, l'extrême bord de l'anneau extérieur est à 130 000 kilomètres du centre, ce qui veut dire que tout le système des anneaux est à l'intérieur de la limite de Roche. On constate la même chose pour l'anneau de Jupiter.

Apparemment, les anneaux de Saturne sont formés de débris qui n'ont jamais pu se rassembler pour former un satellite – comme l'auraient fait des débris situés en dehors de la limite de Roche – ou alors les morceaux d'un satellite qui s'est aventuré trop près, pour une raison ou une autre, et a été réduit en pièces. L'effet de marée diminue avec la taille de l'objet qui le subit, et quand les morceaux descendent en dessous d'une certaine taille, ils ne sont plus affectés, sauf par les chocs éventuels entre eux. D'après certaines estimations, si tous les matériaux des anneaux étaient rassemblés en boule, celle-ci serait un peu plus grande que notre Lune.

LES SATELLITES DE SATURNE

En plus de ses anneaux, comme Jupiter, Saturne possède toute une famille de satellites. Le premier fut découvert par Huygens, en 1656, l'année même où il donna pour la première fois une description de l'anneau. Deux siècles plus tard, ce satellite reçut le nom de Titan (tiré lui aussi de la mythologie grecque). C'est un corps volumineux, presque – mais pas tout à fait – aussi gros que Ganymède. Comme, de plus, il est moins dense, la différence est encore plus nette au point de vue de la masse. Mais Titan est quand même le deuxième satellite du système solaire, que l'on retienne comme critère le volume ou la masse.

Il est néanmoins un domaine où Titan, jusqu'ici, détient le record. Plus éloigné du Soleil, et par conséquent plus froid, que les satellites de Jupiter, il est plus capable de retenir prisonnières les molécules d'un gaz, moins agitées à cette basse température, malgré la valeur relativement faible de la gravité à sa surface. En 1944, un astronome américano-hollandais, Gerard Pieter Kuiper, y a détecté une atmosphère de méthane. (La molécule de méthane est formée d'un atome de carbone et quatre atomes d'hydrogène. C'est le principal constituant du gaz naturel sur la Terre.)

A l'époque de la découverte de Titan, on connaissait en tout cinq satellites : la Lune et les quatre satellites galiléens de Jupiter. Ils avaient tous des tailles voisines, beaucoup plus voisines que celles des différentes planètes. Entre 1671 et 1684, toutefois, Cassini devait découvrir quatre nouveaux satellites de Saturne, tous nettement plus petits qu'Europe, le plus petit des galiléens. Leurs diamètres allaient de 1 500 kilomètres (Japet) à 950 kilomètres (Téthys). Dès lors, il fallait admettre que les satellites pouvaient être petits.

A la fin du XIXe siècle, on connaissait neuf satellites de Saturne. Le dernier en date était Phoebe, découvert par l'astronome américain William Henry Pickering. C'est le plus extérieur des satellites, situé à une distance moyenne de Saturne égale à 13 millions de kilomètres. Il tourne autour de la planète en 549 jours, dans le sens rétrograde. Avec un diamètre de 190 kilomètres, c'est aussi le plus petit des satellites, d'où sa découverte tardive : qui dit petit dit peu lumineux.

Entre 1979 et 1981, trois sondes qui avaient auparavant frôlé Jupiter – *Pioneer 11*, *Voyager 1* et *Voyager 2* – ont photographié en gros plan Saturne, ses anneaux et ses satellites.

Le premier objectif était, bien entendu, Titan, à cause de son atmosphère. Quand les signaux radio de *Voyager 1*, sur leur chemin vers la Terre, traversèrent l'atmosphère de Titan, on parvint à analyser la manière dont cette atmosphère les avait absorbés, et on constata que cette atmosphère

était beaucoup plus dense que prévue. A partir de la quantité de méthane qu'on y avait détectée depuis la Terre, on croyait cette atmosphère aussi ténue que celle de Mars. Or, elle est 150 fois plus dense –1,5 fois plus dense que la nôtre.

C'est que le méthane, seul composant détecté depuis la Terre, ne forme que 2 % de l'atmosphère de Titan, le reste étant de l'azote, gaz difficile à détecter de loin.

Cette atmosphère épaisse est chargée de brouillard, qui cache complètement le sol. Mais ce brouillard est fort intéressant. Le méthane se *polymérise* facilement (c'est-à-dire que ses molécules se combinent volontiers entre elles pour former des molécules plus compliquées). Aussi n'est-il pas impossible que Titan possède des océans de corps organiques assez compliqués. On peut même rêver d'un monde revêtu d'asphalte, avec des collines d'essence solide, et des lacs limpides de méthane et d'éthylène...

Les autres satellites sont, comme on pouvait s'y attendre, criblés de cratères. Mimas (le plus intérieur des neuf) en a même un si grand – relativement à la taille du satellite – que le choc nécessaire pour le produire a presque dû faire voler en éclats le satellite tout entier.

Le deuxième des neuf, Encelade, est relativement lisse, et a peut-être été particulièrement fondu par échauffement dû à l'effet de marée. Hypérion est le plus irrégulier : ses dimensions varient de 110 à 190 kilomètres. Il ressemble assez aux satellites de Mars – en beaucoup plus gros, évidemment. Il est même si gros que, d'après les théories en vigueur, son champ gravitationnel aurait dû l'arrondir en sphère. Peut-être est-il le résultat d'une fracture relativement récente.

Japet, depuis sa découverte en 1671, pose un problème curieux. Il est cinq fois plus brillant lorsqu'il se trouve à l'ouest de Saturne que lorsqu'il se trouve à l'est. Comme il tourne toujours la même face vers la planète, il nous montre tantôt l'un, tantôt l'autre de ses deux hémisphères. On a donc conclu que l'un de ces hémisphères renvoie cinq fois plus de lumière que l'autre, idée confirmée par les photos prises par *Voyager 1*. Un côté de ce satellite est clair et l'autre foncé, comme si le premier était couvert de glace et l'autre de poussières sombres. On ignore les raisons de cette différence.

Les sondes ont découvert huit autres satellites de Saturne, trop petits pour être détectables depuis la Terre, ce qui porte l'effectif total à 17. De ces huit nouveaux satellites, cinq sont plus intérieurs que Mimas, le plus proche se trouvant seulement à 140 000 kilomètres de Saturne (c'est-à-dire à 80 000 kilomètres seulement des nuages qui enveloppent la planète). Il en fait le tour en 14,43 heures seulement.

Deux des satellites, situés juste à l'intérieur de l'orbite de Mimas, présentent cette particularité d'avoir la même orbite, courant sans cesse l'un derrière l'autre tout autour de Saturne. C'est le premier exemple connu d'un tel phénomène. Ils sont à 150 000 kilomètres du centre de Saturne, et leur période de révolution vaut 16,68 heures. En 1967, un astronome français, Audoin Dollfus, a décrit un satellite intérieur à l'orbite de Japet, et l'a baptisé Janus. Or, les résultats qui lui avaient permis de calculer la trajectoire étaient sans doute un mélange de positions des deux satellites précédents, ce qui fait que la trajectoire était fausse. Janus a depuis disparu de la liste des satellites de Saturne.

Les trois autres satellites découverts récemment donnent aussi un exemple d'une situation sans précédent. En effet, on a constaté que Dione,

l'un des satellites découverts par Cassini, partageait son orbite avec un compagnon minuscule. Dione a un diamètre de 1 100 kilomètres, et son compagnon (Dione B), de 30 kilomètres seulement. Sur leur orbite commune, Dione B est toujours de 60 ° en avance sur Dione, autrement dit Saturne, Dione et Dione B forment toujours les sommets d'un triangle équilatéral. C'est ce qu'on appelle une configuration « troyenne », pour des raisons que nous verrons plus loin.

Cette configuration, seulement possible quand le troisième corps est beaucoup plus petit que les deux autres, peut se présenter quand le petit corps est 60 ° devant ou 60 ° derrière l'autre. Dans le premier cas, on parle de position L4, dans le second de position L5 (L est l'initiale de Lagrange, l'astronome français qui a montré, en 1772, qu'une telle configuration était stable).

Mieux encore : Téthys, un autre des satellites de Cassini, a deux compagnons : Téthys B dans la position L4, et Téthys C dans la positon L5 !

De toute évidence, la famille de Saturne est la plus compliquée de celles actuellement connues dans le système solaire.

Quant aux anneaux, les sondes révèlent qu'eux aussi sont beaucoup plus compliqués qu'on le pensait. Vus de près, ils sont formés de centaines – voire de milliers – d'anneaux très minces, l'ensemble évoquant les sillons d'un disque. Par endroits, des « rayons » sombres traversent ce système à angle droit. Enfin, un faible anneau extérieur semble formé de trois anneaux entrelacés.

Tout cela reste pour l'instant inexplicable, mais on pense généralement que des effets électriques viennent se superposer à la gravitation.

Les planètes extérieures

Jusqu'à l'invention du télescope, Saturne resta la plus lointaine des planètes connues, et celle qui se déplaçait le plus doucement. C'était aussi la moins brillante, mais il s'agissait quand même d'un objet de première grandeur. Or, pendant des milliers d'années après la découverte de l'existence des planètes, personne n'a semblé se demander s'il ne pouvait pas en exister qui soient trop lointaines, c'est-à-dire trop peu brillantes, pour être visibles à l'œil nu.

URANUS

Même après que Galilée eut annoncé l'existence de myriades d'étoiles trop discrètes pour être vues sans lunette, la possibilité que des planètes soient dans le même cas ne semblait pas avoir été envisagée.

Et voilà que, le 13 mars 1781, William Herschel (qui n'était pas encore célèbre), au cours de mesures de positions d'étoiles, remarqua dans la constellation des Gémeaux un objet qui était autre chose qu'un point lumineux : un petit cercle. Il supposa d'abord qu'il s'agissait d'une comète – les seuls objets, à part les planètes, à être autre chose que des points. Mais les comètes sont floues, alors que cet objet avait des bords nets. De

plus, il se déplaçait par rapport aux étoiles plus lentement que Saturne, et par conséquent, il devait être plus loin. C'était une planète très éloignée, beaucoup plus éloignée que Saturne – et beaucoup moins brillante. On la baptisa Uranus, du nom grec qui désigne le ciel, et aussi le père de Saturne dans la mythologie.

Située en moyenne à 2,9 milliards de kilomètres du Soleil, Uranus est ainsi deux fois plus loin que Saturne. De plus, c'est une planète plus petite : 52 000 kilomètres de diamètre. C'est encore quatre fois plus que la Terre, et Uranus est aussi une planète géante gazeuse, comme Jupiter et Saturne – mais nettement plus petite que les deux premières. Sa masse vaut 14,5 fois celle de la Terre, mais seulement 1/6,6 celle de Saturne, et 1/22 celle de Jupiter.

A cause de son éloignement et de sa relative petitesse, Uranus est beaucoup moins brillante que Jupiter ou Saturne. Toutefois, cette planète n'est pas vraiment invisible à l'œil nu. Si on regarde au bon endroit par une nuit bien noire, on voit Uranus comme une étoile faible, sans avoir besoin de lunette.

Il semble donc étonnant que les astronomes ne l'aient pas découverte plus tôt. Ils l'on sûrement vue, mais un point aussi peu lumineux n'attire pas l'attention, surtout quand on est persuadé qu'une planète doit nécessairement être brillante. Et même si on l'observe plusieurs nuits de suite, son mouvement est si lent que ses changements de position peuvent bien passer inaperçus. Enfin, les premiers instruments n'étaient pas très bons, et même pointés dans la bonne direction, ne révélaient pas nettement le disque d'Uranus.

Pourtant, dès 1690, l'astronome anglais John Flamsteed avait catalogué une étoile dans la constellation du Taureau, sous le nom de 34 Tauri. Aucun astronome ne parvint à observer à nouveau cette étoile, mais après la découverte d'Uranus et la détermination précise de son orbite, on s'aperçut qu'elle passait bien par la position assignée par Flamsteed à 34 Tauri... De même, un demi-siècle plus tard, l'astronome français Pierre Charles Le Monnier vit Uranus à treize reprises, et crut avoir repéré treize étoiles différentes...

Les estimations varient quant à la période de rotation : on donne habituellement la valeur de 10,82 heures, mais en 1977 on proposa celle de 25 heures. On n'aura probablement pas de certitude à ce sujet avant que les sondes ne s'approchent d'Uranus.

En tout cas, une chose dont on est sûr à propos de la rotation d'Uranus, c'est l'inclinaison de son axe : 98 °, c'est-à-dire un peu plus d'un angle droit. Ainsi, pendant qu'elle tourne autour du Soleil en 84 ans, Uranus semble rouler le long de son orbite. Et chaque pôle est plongé alternativement dans un jour continu de quarante-deux ans, puis une nuit continue de même durée.

Compte tenu de la distance à laquelle Uranus se trouve du Soleil, cela n'a pas grande importance. Mais si la Terre tournait ainsi, les saisons seraient si tranchées que la vie aurait eu beaucoup de mal à s'y développer.

Après avoir découvert Uranus, Herschel continua à l'observer de temps en temps, et en 1787 il lui trouva deux satellites, qui furent baptisés Titania et Obéron. En 1851, l'astronome anglais William Lassell en découvrit deux de plus, plus proches de la planète, Ariel et Umbriel. Enfin, en 1948, Kuiper découvrit Miranda, encore plus proche d'Uranus.

Tous ces satellites tournent autour d'Uranus dans son plan équatorial, de sorte que non seulement la planète, mais tout le système, semble avoir basculé à 98 ° par rapport à l'orbite décrite autour du Soleil.

Les satellites d'Uranus sont très proches de la planète. Il n'y a pas de satellites éloignés – tout au moins, on n'en a pas détecté. Le plus lointain des cinq est Obéron, à 590 000 kilomètres du centre de la planète (une fois et demie la distance Terre-Lune). Le plus rapproché, Miranda, n'est qu'à 130 000 kilomètres du centre d'Uranus.

Aucun de ces satellites n'a une taille de l'ordre de celle des satellites galiléens de Jupiter, de Titan ou de la Lune. Le plus grand, Obéron, n'a que 1 600 kilomètres de diamètre, et le plus petit, Miranda, 220 kilomètres seulement.

Rien ne semblait rendre ce système particulièrement intéressant, quand en 1973 un astronome britannique, Gordon Tayler, montra qu'Uranus devait passer devant une étoile de neuvième grandeur, SAO158687. L'événement intéressait les astronomes, parce que la lumière de l'étoile allait traverser l'atmosphère éventuelle de la planète, et permettre d'étudier celle-ci, aussi bien juste avant le passage que juste après. Les modifications de cette lumière pouvaient donner des indications sur la température, la pression et la composition de l'atmosphère d'Uranus. Cette *occultation* était prévue pour le 10 mars 1977. Cette nuit-là, une équipe d'astronomes américains, dirigée par James L. Elliot, prit place dans un avion qui l'emmena en haute altitude, au-dessus de l'absorption et des distorsions de la basse atmosphère terrestre.

Juste avant qu'Uranus atteigne l'étoile, la lumière de celle-ci baissa brusquement pendant sept secondes, puis augmenta à nouveau. Uranus continua à s'approcher, et quatre autres « baisses d'intensité » se produisirent, durant environ une seconde chacune. Quand l'étoile émergea de l'autre côté, les mêmes épisodes se déroulèrent dans l'ordre inverse. La seule explication possible était l'existence autour d'Uranus d'un système d'anneaux, trop ténus et trop peu lumineux pour être visibles depuis la Terre.

Des observations soigneuses, lors de l'occultation par Uranus d'autres étoiles, notamment le 10 avril 1978, ont révélé un total de neuf anneaux, s'étendant entre 40 000 et 49 000 kilomètres du centre de la planète –bien en dessous de la limite de Roche.

On a pu calculer que ces anneaux étaient si ténus, formés d'une matière si dispersée, qu'ils étaient trois millions de fois moins brillants que ceux de Saturne. Il n'est donc pas surprenant qu'ils ne soient détectables que par des méthodes indirectes comme celle-là.

Plus tard, lors de la découverte de l'anneau de Jupiter, on a commencé à penser que peut-être toutes les planètes géantes gazeuses avaient un anneau, en plus de leurs nombreux satellites. Ce qui caractérise Saturne, c'est seulement la taille et l'éclat de son système d'anneaux.

NEPTUNE

Peu de temps après la découverte d'Uranus, on put calculer son orbite. Mais au fil des années, on constata que la planète ne suivait pas exactement l'orbite calculée. En 1821 un astronome français, Alexis Bouvard, calcula à nouveau cette orbite, en tenant compte de toutes les observations, même celles de Flamsteed. Uranus ne suivait pas mieux cette nouvelle orbite...

Or, dans le calcul de ces orbites, on avait tenu compte des *perturbations* apportées par la faible attraction gravitationnelle des autres planètes, qui

s'exerce sur Uranus et tantôt la ralentit, tantôt l'accélère légèrement. Si l'orbite n'était pas correcte, c'était à coup sûr qu'une autre planète, inconnue et plus lointaine, exerçait elle aussi des perturbations.

A partir des écarts d'Uranus on put calculer la position et les caractéristiques de cette planète. C'était un calcul extrêmement long et ardu, et deux hommes allaient s'y attaquer en même temps, à l'insu l'un de l'autre.

Le premier était John Couch Adams, étudiant en mathématiques à l'université de Cambridge. En 1841, âgé de vingt-deux ans, il commença le calcul, auquel il allait consacrer tous ses loisirs. Quatre ans plus tard, son calcul était fini : il avait déterminé l'endroit où devait se trouver la planète inconnue. Malheureusement, il ne parvint pas à intéresser suffisamment les astronomes anglais pour qu'ils pointent leurs instruments dans cette direction.

Pendant ce temps, un jeune astronome français, Urbain Jean Joseph Le Verrier, qui ignorait complètement l'existence même d'Adams, calculait lui aussi la position de la nouvelle planète. Plus heureux que son concurrent inconnu, Le Verrier s'adressa pour les vérifications à un astronome allemand, Johann Gottfried Galle, et lui demanda de fouiller une certaine région du ciel à la recherche de la planète. Galle possédait justement une carte récente de la région correspondante. Il ne lui fallut même pas une heure d'observation, avec son assistant Heinrich Ludwig D'Arrest, pour découvrir un objet de huitième magnitude qui ne figurait pas sur la carte.

C'était la planète ! Et elle se trouvait, à très peu de choses près, à l'endroit exact indiqué par Le Verrier. On la baptisa Neptune, du nom du dieu de la mer, à cause de sa couleur verdâtre. Et le malheureux Adams partagea désormais avec Le Verrier les honneurs de la découverte...

Neptune se trouve à 4,5 milliards de kilomètres du Soleil (une fois et demie plus loin qu'Uranus, et trente fois plus loin que la Terre). Pour accomplir une révolution autour du Soleil, il lui faut 164,8 années.

C'est presque une « sœur jumelle » d'Uranus (au sens où Vénus est jumelle de la Terre), au moins en ce qui concerne les dimensions. Son diamètre (50 000 kilomètres) est un peu inférieur, mais Neptune est plus dense, et sa masse est supérieure de 18 % à celle d'Uranus. 17,2 fois plus massive que la Terre, c'est la quatrième géante gazeuse du système solaire.

Le 10 octobre 1846, moins de trois semaines après la découverte de Neptune, on lui découvrit un satellite, baptisé Triton (fils de Neptune dans la mythologie grecque). Il s'agissait encore d'un gros satellite, à peu près aussi massif que Titan. C'était le septième de ce type, et le premier découvert depuis Titan, près de deux siècles plus tôt.

Son diamètre est de 3 900 kilomètres (un peu plus que celui de la Lune). Il se trouve à 360 000 kilomètres du centre de Neptune (presque la distance Terre-Lune). Mais à cause du champ gravitationnel plus fort de Neptune, Triton ne met que 5,88 jours à faire un tour – à peu près cinq fois moins que notre Lune.

Triton tourne autour de Neptune dans le sens rétrograde. Ce n'est pas le seul satellite du système solaire à le faire. Mais les autres (les quatre satellites externes de Jupiter et le plus extérieur de ceux de Saturne) sont très petits, et très éloignés de leur planète. Triton est gros, et rapproché de Neptune. Pourquoi donc se déplace-t-il dans le sens rétrograde ? On n'en sait rien.

Pendant plus d'un siècle, Triton resta le seul satellite de Neptune. Puis, en 1949, Kuiper (qui avait découvert Miranda l'année précédente) détecta un objet très petit et très peu lumineux dans le voisinage de Neptune. C'était un deuxième satellite, auquel on donna le nom de Néréide (porté par les nymphes de la mer dans la mythologie grecque).

Son diamètre vaut à peu près 240 kilomètres, et son mouvement est dans le sens direct. Mais de tous les satellites connus, c'est celui qui a la trajectoire la plus excentrique : il s'approche à 1 400 000 kilomètres de Neptune, et s'en éloigne jusqu'à près de dix millions de kilomètres. Autrement dit, à un bout de son orbite, il est sept fois plus loin de Neptune qu'à l'autre bout. Sa durée de révolution vaut 365,21 jours, soit une année terrestre moins quarante-cinq minutes.

Aucune sonde n'ayant encore exploré Neptune, il n'est pas surprenant que nous ne connaissions pas d'autres satellites, ni un éventuel système d'anneaux. Nous ne savons même pas si Triton a une atmosphère, ce qui est très possible, puisque Titan en a une.

PLUTON

La position et la masse de Neptune rendaient compte presque complètement des « fantaisies » d'Uranus. Pour expliquer les petites différences qui subsistaient, certains astronomes ont pensé qu'il fallait rechercher une autre planète inconnue, encore plus éloignée que Neptune. Le plus acharné dans ses calculs et ses recherches fut Lowell (célèbre pour ses idées sur les canaux martiens).

Cette recherche n'était pas facile. Une planète plus éloignée que Neptune était forcément très peu lumineuse, et se perdait dans la foule innombrable des étoiles faibles. De plus, son mouvement devait être si lent que ses changements de position étaient difficiles à déceler. Quand il mourut en 1916, Lowell n'avait pas trouvé sa planète.

Mais à l'observatoire Lowell, en Arizona, les astronomes continuèrent à chercher. En 1929, un jeune astronome, Clyde William Tombaugh, commença à utiliser pour cela un nouveau télescope, capable de photographier avec une grande netteté une portion relativement importante du ciel.

Il utilisa aussi un *comparateur à clignotement*, qui permettait de voir à tour de rôle, et rapidement, les photos de la même région du ciel prises à quelques jours d'intervalle. En ajustant avec soin les deux plaques photographiques, les étoiles occupent exactement les mêmes places sur les deux, et ne bougent pas du tout quand une image remplace l'autre. Par contre, un corps en mouvement n'est pas au même endroit sur les deux photos, et il saute alternativement d'une position à l'autre quand on alterne les deux vues : il clignote.

Même ainsi, la recherche n'était pas facile : il y avait des dizaines de milliers d'étoiles faibles sur chaque plaque, et il fallait examiner la plaque en détail sur toute sa surface pour s'assurer que parmi cette myriade de points il n'y en avait pas un qui clignotait.

A quatre heures du matin le 18 février 1930, Tombaugh trouva un « clignotant » dans la constellation des Gémeaux. Il suivit le déplacement de l'objet pendant un mois, et annonça le 13 mars qu'il avait trouvé la planète cherchée. On lui donna le nom de Pluton, d'après le dieu des enfers

(parce qu'elle est vraiment loin du Soleil) et aussi parce que les deux premières lettres de ce nom sont les initiales de Percival Lowell.

On calcula l'orbite de Pluton, qui se révéla pleine de surprises. D'abord, cette orbite n'était pas aussi loin du Soleil que l'avaient pensé Lowell et ses collègues. La distance moyenne de Pluton au Soleil n'était que de 5,9 milliards de kilomètres, 30 % de plus que pour Neptune.

Ensuite, cette orbite était beaucoup plus excentrée que celles des autres planètes. Son aphélie était à 7,4 milliards de kilomètres du Soleil, et son périhélie à 4,3 milliards de kilomètres seulement.

Au périhélie, quand Pluton est le plus rapprochée du Soleil, elle en est même plus proche que Neptune, de 100 millions de kilomètres environ ! Pluton fait le tour du Soleil en 248 ans, et pendant chaque révolution il y a une vingtaine d'années pendant lesquels la planète est plus proche du Soleil que Neptune. Il se trouve que les vingt dernières années de ce siècle correspondent à l'une de ces périodes : en ce moment, Pluton est plus proche du Soleil que Neptune.

Il ne faut pas en conclure que les deux orbites se croisent, car celle de Pluton est fortement inclinée par rapport à celle des autres planètes. Elle est inclinée de 17,2 ° par rapport à celle de la Terre, alors que celle de Neptune, par exemple, a une inclinaison négligeable par rapport à la nôtre. Aussi, quand Pluton et Neptune sont à la même distance du Soleil, l'une est située très au-dessous de l'autre : les deux planètes ne s'approchent jamais à moins de 2,4 milliards de kilomètres.

Mais la plus grande surprise réservée par Pluton fut la faiblesse de son éclat, qui indiqua immédiatement que ce n'était pas une planète géante gazeuse. Si sa taille était de l'ordre de celle d'Uranus et de Neptune, elle serait considérablement plus brillante. On a d'abord pensé que Pluton était à peu près de la taille de la Terre.

Mais cette estimation était exagérée. En 1950, Kuiper est parvenu à voir Pluton comme un petit disque qu'il a mesuré, et il en a déduit le diamètre de la planète : 5 800 kilomètres, un peu moins que celui de Mars. Quelques astronomes hésitaient à accepter cette valeur, mais le 28 avril 1965, Pluton est passée très près d'une étoile faible, sans l'occulter comme cela aurait été le cas si Pluton avait eu un diamètre supérieur à l'estimation de Kuiper.

Donc Pluton était trop petite pour influencer sensiblement Uranus. Si une planète distante était responsable des derniers écarts de l'orbite d'Uranus, ce n'était certainement pas Pluton.

En 1955, on s'est aperçu que l'éclat de Pluton variait régulièrement, avec une période de 6,4 jours. On a donc supposé d'abord que Pluton tournait sur elle-même en 6,4 jours, ce qui est beaucoup (Vénus et Mercure ont des périodes beaucoup plus longues, bien sûr, mais c'est à cause du voisinage du Soleil). A quoi cela pouvait-il être dû ?

Le 22 juin 1978, la question était résolue : en examinant des photos de Pluton, l'astronome américain James W. Christy lui a trouvé une forme bossue. Il a examiné d'autres photos, et a conclu finalement que Pluton avait un satellite très proche, à 20 000 kilomètres (de centre à centre). A la distance où nous sommes de Pluton, il n'est pas étonnant que ce satellite ait été difficile à découvrir. Christy l'a baptisé Charon, du nom du passeur des enfers dans la mythologie grecque, chargé de faire traverser le Styx aux défunts qui arrivent dans le royaume souterrain de Pluton.

Charon tourne autour de Pluton en 6,4 jours, ce qui est tout juste le temps mis par Pluton pour tourner sur son axe. Ce n'est pas une

coïncidence. Les deux corps se sont mutuellement ralentis, par effet de marée, jusqu'à ce que chacun tourne constamment la même face vers l'autre. Ils tournent maintenant autour du centre de masse de leur système, rigidement liés l'un à l'autre comme les deux boules d'une haltère par l'attraction gravitationnelle.

C'est le seul couple planète-satellite qui se comporte de cette façon. Ainsi, dans le cas du couple Terre-Lune, la Lune présente bien toujours la même face à la Terre, mais celle-ci n'a pas encore été ralentie au point de tourner toujours la même face vers la Lune, parce qu'elle est beaucoup plus massive, et beaucoup plus difficile à freiner. Si la Lune et la Terre avaient des masses plus voisines, peut-être le système se comporterait-il, lui aussi, comme une haltère.

A partir de la distance entre les deux corps et du temps de rotation du système, on peut calculer sa masse totale : elle vaut seulement un huitième de celle de la Lune. Pluton est encore beaucoup plus petite que prévu !

A partir de l'éclat comparé des deux corps, on peut estimer leur taille relative. Finalement, Pluton semble avoir un diamètre de 3 000 kilomètres (comme Europe, le plus petit des sept gros satellites du système). Charon, lui, doit avoir un diamètre de 1 200 kilomètres, à peu près la taille de Dione, l'un des satellites de Saturne.

Ainsi, les deux objets ont des tailles assez voisines. Pluton n'a probablement pas plus de dix fois la masse de Charon, alors que la Terre a 81 fois celle de la Lune. C'est ce qui explique que le couple Pluton-Charon tourne « en haltère » et pas le couple Terre-Lune. Dans tout le système solaire, le couple Pluton-Charon est ce qui ressemble le plus à une « planète double ». Jusqu'à 1978, le meilleur exemple connu était le couple Terre-Lune.

Les astéroïdes

AU-DELÀ DE L'ORBITE DE MARS

A une exception près, chaque planète est entre 1,3 et 2 fois plus loin du Soleil que la précédente. La seule exception est Jupiter, la cinquième planète : elle est 3,4 fois plus éloignée du Soleil que Mars, la quatrième.

Cette exception préoccupa surtout les astronomes après la découverte d'Uranus, quand on s'est brusquement avisé qu'il pouvait rester des planètes à découvrir. Pouvait-il y en avoir une dans cet intervalle anormal – la planète numéro 4 1/2, en somme – qui aurait réussi à rester ignorée jusque-là ? Un astronome allemand, Heinrich W. M. Olbers, prit la direction d'une équipe qui allait fouiller systématiquement le ciel à la recherche de cette planète.

Pendant que l'équipe achevait ses préparatifs, un astronome italien, Giuseppe Piazzi, qui observait le ciel sans penser à des planètes nouvelles, tomba par hasard sur un objet qui se déplaçait par rapport aux étoiles de jour en jour. Sa vitesse permettait de le placer entre Mars et Jupiter, et son éclat était si faible qu'il devait être très petit. Cette découverte a eu lieu le 1er janvier 1801, le premier jour du XIXe siècle.

A partir des observations de Piazzi, le mathématicien allemand Johann K. F. Gauss a calculé l'orbite de l'objet. Il s'agissait bien d'une nouvelle planète, dont l'orbite s'intercalait entre celle de Mars et celle de Jupiter, exactement là où « il en manquait une » pour que la répartition des planètes soit régulière. Piazzi, qui travaillait en Sicile, donna à sa planète le nom de Cérès, d'après la déesse romaine des moissons, particulièrement honorée dans cette île pendant l'Antiquité.

Sa distance et son faible éclat permirent de calculer que Cérès était vraiment très petite, beaucoup plus petite que n'importe quelle autre planète (la valeur moderne de son diamètre est 1 000 kilomètres). Sa masse ne dépasse probablement pas le quinzième de celle de la Lune, et la planète est nettement plus petite que n'importe lequel des gros satellites du système solaire.

Il semblait improbable que Cérès occupât à elle toute seule l'intervalle entre Mars et Jupiter, et Olbers continua ses recherches malgré la découverte de Piazzi. Effectivement, en 1807, on découvrit dans la région trois planètes de plus, nommées Pallas, Junon et Vesta, toutes plus petites que Cérès, la plus petite, Junon, ayant seulement 100 kilomètres de diamètre.

Ces nouvelles planètes sont si petites que même les meilleurs télescopes de l'époque ne les montrèrent que comme des points lumineux, comme les étoiles. Pour cette raison, Herschel suggéra de les appeler *astéroïdes* (« en forme d'étoile »), et ce nom leur est resté.

Le cinquième n'a été découvert qu'en 1845, par l'astronome allemand Karl L. Hencke, qui l'a baptisé Astrée. Mais ensuite, les découvertes se sont multipliées. A l'heure actuelle, on connaît 1 600 astéroïdes, tous considérablement plus petits que Cérès, la première découverte. Sans aucun doute, il en reste des milliers à découvrir. Presque tous sont situés dans l'intervalle entre Mars et Jupiter, dans la région connue maintenant sous le nom de *ceinture des astéroïdes*.

Quelle peut être la cause de leur existence ? Très tôt, au moment où on n'en connaissait que quatre, Olbers a suggéré que c'étaient les débris d'une planète qui avait explosé. Mais la plupart des astronomes pensaient plutôt qu'il s'agissait des morceaux d'une planète qui ne s'était pas formée. Alors que, dans les autres régions du système solaire, la matière de la nébuleuse primitive s'est condensée en donnant d'abord des planéticules (objets fort semblables aux astéroïdes) qui se sont ensuite rassemblés en planètes (les derniers arrivés laissant leur trace sous forme de cratères), dans la ceinture des astéroïdes la condensation n'aurait pas dépassé le stade des planéticules. Peut-être est-ce à cause du voisinage de Jupiter, le géant du système ?

Un argument en faveur de cette hypothèse est celui des *intervalles de Kirkwood*. En 1866, on connaissait assez d'astéroïdes pour étudier leur répartition dans l'intervalle Mars-Jupiter, et s'apercevoir qu'elle n'était pas régulière. Il y avait des régions où il n'y en avait pas du tout. Aucun astéroïde n'avait une distance moyenne au Soleil voisine de 370 millions de kilomètres, ni de 450 millions, ni de 490, ni de 540.

Un astronome américain, Daniel Kirkwood, a suggéré en 1866 que, sur ces orbites, les astéroïdes avaient une période de révolution qui était une fraction simple de celle de Jupiter. Dans ces conditions, les perturbations dues à Jupiter étaient très importantes, obligeant l'astéroïde à s'éloigner ou à s'approcher davantage. Ces anneaux de Kirkwood montraient bien

que l'influence de Jupiter dans la région était très importante, et permettait peut-être d'expliquer la non-condensation en planète de l'ensemble des astéroïdes.

Un peu plus tard, on a découvert une autre relation entre Jupiter et les astéroïdes. En 1906, un astronome allemand, Max Wolf, a découvert l'astéroïde n° 588. Sa vitesse était notamment faible ; il était donc assez loin du Soleil. C'était en fait le plus lointain découvert jusque-là. On lui a donné le nom d'Achille, le héros grec de la guerre de Troie. (Bien qu'on donne d'habitude des noms féminins aux astéroïdes, on en donne des masculins à ceux qui ont des orbites bizarres.)

On a observé soigneusement Achille, et on a constaté qu'il suivait l'orbite de Jupiter, avec 60 ° d'avance sur la planète géante. Moins d'un an plus tard, on a découvert l'astéroïde n° 617, qui suivait la même orbite avec 60 ° de retard sur Jupiter, et on l'a nommé Patrocle, du nom de l'ami d'Achille dans l'*Iliade* d'Homère. En fait, chacun d'eux était accompagné de plusieurs autres, qui reçurent aussi les noms de héros de la guerre de Troie. Ce fut le premier exemple connu vérifiant la stabilité de la configuration de Lagrange, où trois corps occupent les sommets d'un triangle équilatéral, le Soleil, une planète et un corps beaucoup plus petit. Achille et son groupe occupent la position L4, Patrocle et le sien la position L5. Aussi, on appelle ces positions *troyennes* et les astéroïdes qui les occupent *astéroïdes troyens.*

Les satellites les plus extérieurs de Jupiter, qui semblent avoir été capturés à une époque relativement récente, sont peut-être d'anciens astéroïdes troyens.

Le satellite externe de Saturne, Phoebe, et celui de Neptune, Néréide, sont peut-être aussi des satellites capturés, ce qui laisserait croire à l'existence d'astéroïdes au-delà de l'orbite de Jupiter. Mais peut-être aussi ont-ils été arrachés à la ceinture des astéroïdes sous l'effet de perturbations particulièrement intenses, et capturés ensuite par des planètes.

En 1920, par exemple, Baade découvrit l'astéroïde 944, qu'il baptisa Hidalgo. Son aphélie est très extérieur à l'orbite de Jupiter, et sa période de révolution vaut 13,7 ans, plus que celles de Jupiter et trois fois plus que la période moyenne des astéroïdes. Son orbite est très excentrée (son excentricité vaut 0,66), et au périhélie il est seulement à 300 millions de kilomètres du Soleil, en plein dans la ceinture des astéroïdes, alors que son aphélie est à 1 450 000 000 kilomètres du Soleil – aussi loin que Saturne ! Son orbite est inclinée, et il ne court pas le risque d'être capturé, mais un autre astéroïde sur une orbite aussi excentrée pourrait fort bien se rapprocher davantage de Saturne, et finir par être capturé par la planète – ou par une autre encore plus lointaine.

Est-il inconvevable qu'un astéroïde soit envoyé par des perturbations gravitationnelles sur une orbite entièrement extérieure à la ceinture des astéroïdes ? En 1977, l'astronome américain Charles Kowall a détecté un objet très pâle qui se déplaçait trois fois plus lentement que Jupiter : il devait être bien au-delà de l'orbite de la planète géante.

Kowall le suivit pendant quelques jours, pour avoir une idée grossière de son orbite, puis le chercha sur d'anciennes plaques photographiques. Il le retrouva sur trente d'entre elles, la plus ancienne datant de 1895, ce qui lui donnait assez de positions pour calculer l'orbite exacte.

C'est un astéroïde de taille respectable, avec un diamètre de l'ordre de 200 kilomètres. A l'aphélie, il est aussi éloigné qu'Uranus, et au périhélie

aussi éloigné que Saturne. Mais l'inclinaison de son orbite fait qu'il n'approche ni l'une ni l'autre des deux planètes.

Kowall l'a baptisé Chiron, d'après un centaure de la mythologie grecque. Sa période de révolution vaut 50,7 ans. En ce moment, il est près de son aphélie, mais dans une vingtaine d'années il sera deux fois moins loin, et on pourra le voir plus nettement.

EARTH GRAZERS ET APOLLOS

Si des astéroïdes dépassent l'orbite de Jupiter, ne peut-il y en avoir qui franchissent celle de Mars, et s'approchent plus près du Soleil ?

Le premier a été découvert le 13 août 1898 par un astronome allemand, Gustav Witt. C'est l'astéroïde 433, dont la période est seulement de 1,76 an – 44 jours de moins que celle de Mars. Sa distance moyenne au Soleil doit donc être inférieure à celle de Mars. On lui donne le nom d'Eros.

Son orbite est assez excentrée. A l'aphélie, il est bien à l'intérieur de la ceinture des astéroïdes, mais au périhélie il s'approche à 170 millions de kilomètres du Soleil – à peine plus loin que la Terre. Son orbite étant inclinée, il ne s'approche pas si près de la Terre que si les deux orbites étaient dans le même plan.

Mais si Eros et la Terre sont bien placés sur leurs orbites, la distance qui les sépare peut diminuer jusqu'à 22,5 millions de kilomètres, un peu plus de la moitié de la distance minimale entre Vénus et la Terre : au moment de sa découverte, Eros était notre plus proche voisin – à part la Lune.

Ce n'est pas un corps énorme. D'après ses changements d'éclat, il a la forme d'une brique, de dimension moyenne, 15 kilomètres environ. Ce n'est pourtant pas négligeable : si un corps de cette taille heurtait la Terre, ce serait une catastrophe inimaginable.

En 1931, Eros devait s'approcher à seulement 26 millions de kilomètres de la Terre, ce qui allait permettre de calculer avec précision sa parallaxe – et donc toutes les distances du système solaire. Un vaste projet fut mis en place et réussit : les résultats ne furent améliorés que par la réflexion sur Vénus d'un faisceau radar.

Tout astéroïde qui, comme Eros, s'approche plus de la Terre que Vénus reçoit le nom – un peu exagéré – d'*earth grazer* (« frôleur de la Terre »...). Entre 1898 et 1932, on en a découvert seulement trois, dont aucun ne s'approchait aussi près qu'Eros.

Le record est tombé pourtant, le 12 mars 1932, quand un astronome belge, Eugène Delporte, a découvert l'astéroïde 1221 : son orbite était analogue à celle d'Eros, mais s'approchait à 16 millions de kilomètres de celle de la Terre. On lui a donné le nom d'Amor (l'équivalent latin du grec Eros).

Juste six semaines plus tard, le 24 avril 1932, l'astronome allemand Karl Reinmuth a découvert un autre earth grazer, qu'il a nommé Apollo. Cet astéroïde étonnant passe, au périhélie, à moins de 100 millions de kilomètres du Soleil : il pénètre non seulement en deçà de l'orbite de Mars, mais aussi de celle de la Terre, et même de Vénus ! Mais son excentricité est telle qu'à l'aphélie il s'éloigne à 350 millions de kilomètres du Soleil, plus loin encore qu'Eros. Le 15 mai 1932, Apollo s'est approché à 10 millions de kilomètres de la Terre – moins de trente fois la distance de

la Lune. Cet astéroïde mesure un peu plus d'un kilomètre de diamètre, ce qui est tout à fait suffisant pour rendre épouvantable un choc éventuel. Depuis lors, on appelle *apollo* tout astéroïde qui s'approche plus du Soleil que Vénus.

En février 1936, Delporte, qui avait découvert Amor quatre ans plus tôt, a récidivé avec un autre earth grazer, qu'il a baptisé Adonis. Quelques jours avant sa découverte, Adonis était passé à 2,5 millions de kilomètres de la Terre – 6,3 fois la distance de la Lune. Son périhélie est à 60 millions de kilomètres, près de l'orbite de Mercure ! C'est le deuxième apollo.

En novembre 1937, Reinmuth (qui avait découvert Apollo) a annoncé la découverte du troisième objet de ce type, qu'il a baptisé Hermès, et qui est passé à 800 000 kilomètres de la Terre, deux fois la distance de la Lune. Avec le peu d'informations dont il disposait, Reinmuth a calculé une orbite approximative, montrant que, dans le « meilleur » des cas, Hermès pouvait passer à 300 000 kilomètres de la Terre (moins loin que la Lune). Toutefois, on n'a jamais revu Hermès depuis cette date.

Le 26 juin 1949, Baade a découvert un apollo encore plus bizarre. Sa période est de 1,12 année, et son excentricité la plus élevée de toutes celles des astéroïdes connus : 0,827. A l'aphélie, il passe sagement dans la ceinture des astéroïdes, entre Mars et Jupiter. Mais son périhélie est à 29 millions de kilomètres du Soleil – plus près que Mercure ! Baade l'a appelé Icare, le nom du jeune Grec de la mythologie qui se serait élancé dans les airs, muni d'ailes fabriquées par son père Dédale, et se serait approché trop près du Soleil ; celui-ci aurait fait fondre la cire qui maintenait ses plumes et provoqué sa chute...

Depuis 1949, on a découvert d'autres apollos. Certains ont des périodes de moins d'un an, et l'un d'entre eux est plus proche du Soleil que la Terre, en tous les points de son orbite. En 1983, on en a découvert un qui s'approchait plus du Soleil qu'Icare.

Certains astronomes estiment qu'il y a, dans l'espace, environ 750 apollos de diamètre supérieur à 800 mètres. On estime qu'en un million d'années, quatre de ces objets heurtent la Terre, trois heurtent Vénus, un Mercure, Mars et la Lune, et sept autres sont perturbés à tel point qu'ils quittent pour toujours le système solaire. Pourtant leur nombre total ne décroît pas : de nouveaux apollos sont crées par perturbation de certains astéroïdes normaux, situés entre Mars et Jupiter.

Les comètes

D'autres membres du système solaire approchent parfois le Soleil de près. Ils apparaissent à l'œil comme des objets flous, faiblement lumineux, qui s'étendent à travers le ciel – nous l'avons vu au chapitre 2 – comme des étoiles brouillées, munies de longues « queues » ou de « chevelures flottantes ». Pour cette raison, les anciens Grecs leur donnèrent le nom d'*aster kometes* (étoiles chevelues), et on les appelle aujourd'hui des *comètes*.

Contrairement aux étoiles et aux planètes, les comètes ne semblent pas suivre des chemins facilement prévisibles : elles apparaissent et disparaissent sans ordre et sans régularité. Quand les gens croyaient que les étoiles

et les planètes avaient une influence sur la vie des hommes, les allées et venues inopinées des comètes leur semblaient associées à des événements exceptionnels de la vie – des catastrophes soudaines, par exemple.

C'est seulement en 1473 qu'un Européen réagit à l'apparition d'une comète autrement qu'en frissonnant de peur : cette année-là, l'astronome allemand Regiomontanus nota, nuit après nuit, le déplacement d'une comète par rapport aux étoiles.

En 1532, deux astronomes – l'Italien Girolamo Fracastoro et l'Allemand Peter Apian – étudièrent une comète et annoncèrent que sa queue était toujours tournée à l'opposé du Soleil.

En 1577, une autre comète apparut, et Tycho Brahé tenta de déterminer sa distance en mesurant sa parallaxe. Si, comme le disait Aristote, les comètes étaient des phénomènes atmosphériques, cette parallaxe devait être plus grande que celle de la Lune. Or, il n'en était rien : elle était trop petite pour pouvoir être mesurée ! La comète était donc plus éloignée que la Lune : c'était un objet astronomique.

Mais à quoi pouvait bien être due l'irrégularité de l'apparition des comètes ? Quand Newton publia en 1687 sa théorie de la gravitation universelle, il semblait pourtant bien que les comètes, comme les autres objets du système solaire, devaient être soumises au champ de gravitation du système.

Or, en 1682, une comète apparut, et Edmund Halley, un ami de Newton, observa son déplacement à travers le ciel. En fouillant des comptes rendus plus anciens, il constata que les comètes de 1456, 1531 et 1607 avaient suivi un chemin similaire. Ces comètes s'étaient succédé à des intervalles de 75 ou 76 ans.

Halley réalisa que les comètes tournaient autour du Soleil comme les planètes, mais sur des orbites qui étaient des ellipses très allongées. Elles passaient la plus grande partie de leur temps dans les parages extrêmement lointains de leur aphélie, trop distantes et trop faibles pour être visibles, puis passaient très vite à travers leur périhélie. C'est seulement à ce moment qu'on les voyait et, faute de pouvoir observer le reste de leur mouvement, celui-ci semblait aléatoire et capricieux.

Aussi Halley prédit que la comète de 1682 reviendrait en 1758. Il n'était plus là pour la voir, mais elle revint bien, et se montra pour la première fois le 25 décembre 1758 : elle était un peu en retard, parce qu'elle avait été un peu ralentie au passage par l'attraction de Jupiter. Cette comète porte dorénavant le nom de comète de Halley. Elle est réapparue en 1832 et en 1910. Revenue encore une fois en 1986, comme prévu, elle avait été repérée dès 1983 par les astronomes : un objet très lointain et difficile à discerner, mais se rapprochant de la Terre.

Depuis Halley, on a déterminé les orbites de plusieurs autres comètes : ce sont toutes des comètes à période courte, dont les orbites sont entièrement situées dans les limites du système solaire. Ainsi la comète de Halley, à son périhélie, est à 88 millions de kilomètres du Soleil, juste à l'intérieur de l'orbite de Vénus. A l'aphélie, elle se trouve à 5,3 milliards de kilomètres du Soleil, au-delà de l'orbite de Neptune.

L'orbite la plus courte correspond à la comète Encke, qui tourne autour du Soleil en 3,3 ans. Au périhélie elle est à 51 millions de kilomètres du Soleil (comme Mercure), et à l'aphélie à 610 millions de kilomètres, dans la ceinture des astéroïdes : c'est la seule comète connue dont l'orbite soit entièrement à l'intérieur de celle de Jupiter.

Quant aux comètes à longue période, leurs aphélies sont très au-delà des limites du système, et elles ne reviennent vers nous que tous les quelques millions d'années. En 1973, l'astronome tchèque Lajos Kohoutek a découvert une nouvelle comète qui souleva un grand intérêt parce qu'on s'attendait à ce qu'elle fût très brillante. Malheureusement il n'en était rien. Au périhélie, cette comète passe à 38 millions de kilomètres seulement du Soleil – plus près que Mercure. Mais à l'aphélie (si le calcul de son orbite est correct) elle s'éloigne à 500 milliards de kilomètres – 20 fois plus loin que Neptune. Aussi, la comète Kohoutek doit-elle avoir une période de révolution autour du Soleil de l'ordre de 200 000 ans. Il en existe certainement d'autres avec des périodes plus longues encore – et des orbites plus étendues.

En 1950, Oort a estimé que, dans un rayon de quatre à huit mille milliards de kilomètres autour du Soleil, soit 25 fois la distance de la comète Kohoutek à son aphélie, il pouvait exister une centaine de milliards de petits objets, avec des diamètres allant du demi-kilomètre à une dizaine de kilomètres. Rassemblés en boule, leur masse totale vaudrait environ un huitième de celle de la Terre.

Ce sont des restes du nuage originel de poussières et de gaz qui s'est condensé il y a cinq milliards d'années pour donner naissance au système solaire. Les comètes se distinguent des astéroïdes par leur composition : ces derniers sont rocheux, tandis que les comètes sont surtout formées de glaces dures comme la pierre aux distances auxquelles elles se trouvent du Soleil, mais qui s'évaporeraient facilement à la chaleur. L'astronome américain Fred Lawrence Whipple a été le premier à suggérer, en 1949, que les comètes étaient formées de glaces, avec peut-être un noyau rocheux, ou des cailloux disséminés dans la masse : c'est le modèle de « la boule de neige sale ».

Les comètes ordinaires restent toujours dans les régions très éloignées, tournant lentement autour du Soleil avec des périodes de l'ordre du million d'années. Mais de temps en temps, une collision ou une autre perturbation, comme l'influence d'une autre étoile, vient accélérer une comète, qui quitte alors le système solaire, ou au contraire la ralentir, ce qui la fait plonger vers le Soleil, retourner à son point de départ et plonger à nouveau, etc. Ce sont celles-là que nous voyons, quand elles s'approchent suffisamment de la Terre.

Ayant leur origine dans une « coquille » sphérique, les comètes peuvent venir sous n'importe quel angle, et ont autant de chances de se déplacer dans le sens direct que dans le sens rétrograde. La comète de Halley, par exemple, est rétrograde.

Quand la comète s'approche du Soleil, la chaleur vaporise de la glace et libère les poussières qui y étaient enfermées. La vapeur et la poussière forment une sorte de nuage autour de la comète (la *coma*) et la font apparaître comme un gros objet flou.

Ainsi la comète de Halley, quand elle est complètement gelée, a peut-être seulement un diamètre de l'ordre de deux kilomètres. Mais quand elle passe près du Soleil, son nuage peut avoir jusqu'à 400 000 kilomètres de diamètre, occupant vingt fois le volume de Jupiter, la matière qui forme ce nuage étant très dispersée, guère plus qu'une sorte de « vide brumeux »...

Or le soleil émet constamment des particules très petites, plus petites que des atomes (nous en parlerons au chapitre 7), qui filent dans toutes les directions. Ce *vent solaire* vient frapper le nuage de la comète, et le

balaie vers l'extérieur en une longue queue, qui peut être plus volumineuse que le Soleil lui-même, mais où la matière est encore plus ténue que dans le nuage. Bien entendu, cette queue est toujours tournée à l'opposé du Soleil, comme l'avaient remarqué Frascatoro et Apian il y a quatre siècles et demi.

A chaque passage près du Soleil, la comète perd une partie de ses matériaux, qui se vaporisent et s'éloignent dans la queue. Finalement, après quelques centaines de passages, elle se réduit en poussière et disparaît, ou se réduit à un noyau rocheux (ce que semble faire la comète Hencke), qui ne différera pas d'un astéroïde.

Durant la longue histoire du système solaire, des millions de comètes ont sûrement été chassées loin de lui, ou au contraire attirées vers le Soleil, ce qui a fini par les faire disparaître. Mais il en reste de nombreux milliards. Nous ne risquons pas de manquer de comètes.

Chapitre 4

La Terre

Sa forme et sa taille

Un Soleil énorme, quatre planètes géantes, cinq plus petites, une quarantaine de satellites, plus de cent mille astéroïdes et peut-être cent milliards de comètes : voilà le système solaire. Or, pour autant que nous sachions aujourd'hui, la vie existe sur un seul de ces objets, notre Terre. C'est vers elle, par conséquent, que nous allons maintenant nous tourner.

ELLE EST RONDE

C'était l'une des idées les plus importantes chez les anciens Grecs. Elle semble leur être venue d'abord (on l'attribue à Pythagore, vers 525 av. J.-C.) pour des raisons philosophiques – par exemple la perfection de la forme sphérique. Mais les Grecs ne tardèrent pas à vérifier cette forme par l'observation. Vers 350 av. J.-C., Aristote dressa la liste des arguments déjà réunis en faveur de cette idée que la Terre n'était pas plate, mais ronde. Le plus important fut l'apparition, dans un voyage nord-sud, de nouvelles étoiles à l'horizon vers lequel on allait et la disparition d'autres étoiles derrière l'horizon dont on s'éloignait. D'autre part, quand un navire s'éloignait, sa coque disparaissait la première, quelle que fût la direction de son voyage. Enfin, l'ombre de la Terre sur la Lune, lors des éclipses de Lune, était toujours un cercle, quelle que fût à ce moment la position de la Lune. Ces deux derniers arguments prouvaient, sans aucun doute possible, que la Terre était une sphère.

Au moins parmi les érudits, cette idée ne disparut jamais, même pendant la nuit du Moyen Age. Le poète italien Dante Alighieri a décrit une Terre sphérique dans le résumé des idées médiévales que constitue la *Divine Comédie.*

Mais il en allait tout autrement de l'idée de la rotation de la Terre. Dès 350 av. J.-C., le philosophe grec Héraclide du Pont suggéra qu'il serait beaucoup plus simple d'imaginer la rotation de la Terre autour de son axe que celle de l'Univers entier autour d'une Terre immobile. Mais cette idée choqua presque tous les savants de l'Antiquité et du Moyen Age, et jusqu'en 1632 : c'est sur ce point que Galilée fut condamné par l'Inquisition de Rome, et contraint de renier publiquement ses convictions.

Pourtant, la théorie de Copernic rendit complètement illogique l'idée d'une Terre immobile, et peu à peu tout le monde accepta qu'elle tournait. Toutefois, c'est seulement en 1851 que cette rotation fut démontrée de façon expérimentale. Cette année-là, le physicien français Jean Bernard Léon Foucault suspendit un long pendule à la coupole d'une église parisienne. Tous les physiciens étaient d'accord : un tel pendule allait osciller dans un plan fixe, indépendamment de la rotation de la Terre. Au pôle Nord, par exemple, il oscillerait dans un plan fixe, tandis que la Terre tournerait au-dessous, en vingt-quatre heures, d'un tour en sens inverse des aiguilles d'une montre. Pour un observateur, entraîné par la Terre – qui lui semblerait donc immobile –, le pendule semblerait osciller dans un plan tournant, qui décrirait un tour en vingt-quatre heures dans le sens des aiguilles d'une montre. Au pôle Sud, une seule différence : le plan d'oscillation semblerait tourner dans l'autre sens.

A des latitudes plus basses, le plan d'oscillation du pendule devait encore avoir un mouvement apparent de rotation (dans le sens inverse des aiguilles d'une montre dans l'hémisphère sud, et dans l'autre sens dans l'hémisphère nord), mais de plus en plus lentement à mesure que l'on s'éloignerait des pôles. A l'équateur le plan d'oscillation ne tournerait pas du tout.

Lors de l'expérience de Foucault, le plan d'oscillation tourna dans le bon sens, et à la vitesse prévue. En quelque sorte, les spectateurs virent la Terre tourner sous le pendule.

La rotation de la Terre a de nombreuses conséquences. Sa surface se déplace plus vite à l'équateur, où elle doit décrire 40 000 kilomètres en 24 heures, c'est-à-dire faire 1 600 km/h. Quand on s'éloigne de l'équateur, cette vitesse diminue, puisque le cercle parcouru en 24 heures devient de plus en plus petit. Près des pôles, ce cercle devient vraiment très petit, et les pôles eux-mêmes sont immobiles.

L'air participe du mouvement de la surface au-dessus de laquelle il se trouve. Si une masse d'air quitte l'équateur vers le nord, sa vitesse (égale à celle de la surface à l'équateur) est plus grande que celle de la surface qu'il va survoler. Cet air « dépasse » donc la surface dans son mouvement vers l'est, et se déplace – par rapport à elle – vers l'est. Cette dérive est un exemple de l'*effet de Coriolis,* du nom du mathématicien français qui l'a étudié le premier en 1835.

L'effet de Coriolis fait tourner les masses d'air dans le sens des aiguilles d'une montre dans l'hémisphère nord. Dans l'hémisphère sud, l'effet est inversé, et la déviation est dans le sens inverse des aiguilles d'une montre. Dans les deux cas, des *perturbations cycloniques* s'établissent. Les grandes tempêtes de ce type s'appellent *ouragans* dans l'Atlantique nord et *typhons*

dans le Pacifique. Des tempêtes plus limitées mais plus intenses reçoivent les noms de *cyclones* ou de *tornades*. Au-dessus de la mer, ces tourbillons violents suscitent des *trombes d'eau* spectaculaires.

Toutefois, la conséquence la plus intéressante de la rotation de la Terre avait été prévue deux siècles avant l'expérience de Foucault, à l'époque de Newton. L'idée de la perfection sphérique de la Terre était imposée depuis deux mille ans, quand Newton commença à se demander quel effet allait produire sur cette sphère une rotation autour d'un axe. Il nota la différence de vitesse de différents points de la surface de la Terre, à différentes latitudes, et chercha à en évaluer les conséquences.

Plus la rotation est rapide, plus grande est la force centrifuge, qui tend à éloigner de l'axe la matière en rotation. Par conséquent l'effet de cette force, nul aux pôles, doit augmenter à mesure qu'on s'en éloigne, et devenir maximal à l'équateur, qui tourne à toute vitesse. En d'autres termes, le milieu de la Terre doit avoir tendance à s'écarter de l'axe : la Terre doit être un *sphéroïde aplati* avec un renflement équatorial et un aplatissement aux pôles. Elle doit plutôt ressembler à une mandarine qu'à une balle de golf. Newton parvint même à calculer que cet aplatissement polaire devait être à peu près égal à 1/230 du diamètre, résultat étonnamment proche de la réalité.

La Terre tourne si lentement que ces déformations sont trop légères pour être faciles à détecter. Mais déjà à l'époque de Newton, deux observations astronomiques encouragèrent cette idée. Tout d'abord, cet aplatissement est parfaitement visible dans les cas de Jupiter et Saturne, comme nous l'avons vu au chapitre précédent.

Ensuite, si la Terre est vraiment renflée à l'équateur, l'attraction variable exercée sur ce renflement par la Lune, qui est le plus souvent au nord ou au sud de l'équateur, lors de son mouvement autour de la Terre, devrait déplacer l'axe de rotation de la Terre, en lui faisant décrire un cône, chaque pôle pointant ainsi vers un point lentement variable dans le ciel. Ces points décrivent un cercle en 25 750 ans. Hipparque avait d'ailleurs noté ce déplacement en 150 av. J.-C., en comparant les positions des étoiles de son temps et celles qui avaient été notées un siècle et demi plus tôt. A cause de ce mouvement de l'axe de la Terre, la position du Soleil à l'équinoxe se décale chaque année de cinquante secondes d'arc vers l'est. L'équinoxe se produit donc un peu plus tôt chaque année, et Hipparque appela le phénomène *précession des équinoxes*, nom encore utilisé de nos jours.

Bien sûr, le monde scientifique chercha d'autres preuves de l'aplatissement de la Terre. On eut donc recours à une technique traditionnelle en matière de géométrie : la trigonométrie. Sur une surface courbe, la somme des angles d'un triangle est supérieure à 180°. Si la Terre était aplatie, comme le disait Newton, cet excès devait être plus grand au voisinage de l'équateur qu'au voisinage – plus plat – des pôles. Vers 1730, les Français réalisèrent les premières mesures, en dressant des cartes précises de deux régions, dans le Nord et le Sud de la France. Sur la base de ces mesures, l'astronome français Jacques Cassini (fils de celui qui avait noté l'aplatissement de Jupiter et de Saturne) conclut que la Terre devait être allongée au lieu d'être aplatie ! En exagérant beaucoup, sa forme était plutôt celle d'un concombre que d'une mandarine...

Mais la différence de courbure entre le Nord et le Sud de la France était évidemment trop faible pour donner des résultats concluants. Aussi, en

1735 et 1736, deux expéditions françaises partirent vers deux régions de latitudes très différentes, l'une vers le Pérou, près de l'Équateur, et l'autre en Laponie, près des régions arctiques. En 1744, leurs résultats ne laissaient plus de place au doute : la surface de la Terre était plus courbe au Pérou qu'en Laponie.

Les meilleures mesures actuelles montrent que le diamètre équatorial de la Terre mesure 42,5 kilomètres de plus que son diamètre polaire (chacun valant à peu près 12 800 kilomètres).

Ces recherches du XVIII[e] siècle, au sujet de la forme de la Terre, révélèrent à la communauté scientifique l'imperfection des techniques de mesure. Aucun étalon digne de confiance n'existait qui permette des mesures précises. Ce fut l'une des causes de l'adoption par la Révolution française, un demi-siècle plus tard, d'un système d'unités logique et scientifiquement établi, le *système métrique,* fondé sur le mètre. Le système métrique est maintenant utilisé par les scientifiques du monde entier, à leur parfaite satisfaction, et c'est aussi le système usuel partiquement partout.

On ne peut pas attacher trop d'importance à l'existence d'étalons de mesure très précis. Une bonne part de l'effort scientifique vise constamment à améliorer ces étalons. Le mètre étalon et le kilogramme étalon ont été d'abord fabriqués en un alliage de platine et d'iridium pratiquement inaltérable, et gardés dans les environs de Paris avec le plus grand soin – en particulier à température constante, pour éviter la dilatation ou la contraction.

Puis on a découvert de nouveaux alliages comme l'Invar (abréviation d'« invariable »), composé de nickel et de fer dans des proportions convenables, alliages très peu affectés par les variations de température. On a pu donc les utiliser pour fabriquer de meilleurs étalons, et le physicien français d'origine suisse, Charles Edouard Guillaume, inventeur de l'Invar, a reçu pour cela le prix Nobel en 1920.

Pourtant, en 1960, la communauté scientifique renonça aux étalons de longueur solides. La Conférence internationale des poids et mesures a adopté comme étalon la longueur d'onde de l'émission lumineuse d'une variété du krypton. Il faut 1 650 763 ,73 de ces longueurs d'onde – beaucoup plus invariables que n'importe quel objet fabriqué par l'homme – pour faire un mètre, défini ainsi avec une précision mille fois meilleure qu'auparavant. Puis, en 1984, le mètre a été lié à la vitesse de la lumière, et défini comme la distance parcourue par la lumière pendant le nombre convenable de milliardièmes de seconde.

MESURER LE GÉOÏDE

La forme obtenue en égalisant la surface de la Terre partout au niveau de la mer s'appelle *le géoïde.* Bien sûr, en réalité cette surface est irrégulière, marquée de montagnes, de ravins, et ainsi de suite. Même avant que Newton soulève la question de l'aplatissement, les savants avaient déjà essayé d'évaluer ces écarts par rapport à ce qu'ils croyaient être une sphère parfaite. Pour cela, ils avaient utilisé un pendule.

C'est en 1581 que Galilée, alors âgé de dix-sept ans, établit que la période d'un pendule ne dépend que de sa longueur, et pas de l'amplitude de ses oscillations. On raconte qu'il aurait fait cette découverte en observant les balancements des lustres de la cathédrale de Pise, pendant la messe. On

montre encore dans cette cathédrale la *lampe de Galilée,* mais elle date seulement de 1584. Plus tard, Huygens relia un pendule aux rouages d'une horloge, mettant à profit la régularité de son mouvement pour améliorer celle de l'horloge. En 1656, il réalisa ainsi la première *pendule* multipliant par dix la précision des mesures de temps.

La période du pendule dépend à la fois de sa longueur et du champ de gravitation. Au niveau de la mer, un pendule d'un mètre « bat la seconde », c'est-à-dire que sa période vaut deux secondes. Une relation équivalente avait déjà été établie en 1644 par Marin Mersenne, mathématicien français et élève de Galilée. On peut se servir d'un tel pendule pour étudier les variations du champ de gravitation, et donc les irrégularités de la Terre : si sa période d'oscillation vaut deux secondes au niveau de la mer, elle sera un peu plus longue en haut d'une montagne, où le champ de gravitation est plus faible parce que le centre de la Terre est plus éloigné.

En 1673, une expédition française en Guyane, sur la côte nord de l'Amérique du Sud (près de l'équateur), constata qu'à cet endroit, au niveau de la mer, le pendule battait plus lentement. Un peu plus tard, Newton vit là une preuve du renflement équatorial, qui rend le niveau de la mer, à cet endroit, plus éloigné qu'en Europe du centre de la Terre, ce qui diminue le champ de gravitation. Après la double expédition au Pérou et en Laponie qui prouva la justesse de sa théorie, un membre de l'expédition de Laponie, le mathématicien français Alexis Claude Clairaut, élabora des méthodes de calcul de l'aplatissement de la Terre à partir des indications du pendule. On peut ainsi déterminer la forme du géoïde, c'est-à-dire de la Terre « uniformisée au niveau de la mer ». On constate que cette forme s'écarte très peu du sphéroïde aplati parfait (moins de cent mètres d'écart en tous points). De nos jours, on peut aussi mesurer le champ de gravitation avec un *gravimètre,* qui comporte une masse suspendue à un ressort très sensible. L'allongement de ce ressort, placé devant une graduation, indique la force avec laquelle la masse est tirée vers le bas, et donc le champ de gravitation, avec une grande précision.

Le champ de gravitation au niveau de la mer varie environ de 0,6 %, étant bien entendu minimal à l'équateur. Dans la vie de tous les jours, cela ne se remarque pas, mais cela peut affecter les records sportifs. Les résultats aux Jeux olympiques dépendent dans une certaine mesure de la latitude (et de l'altitude) de la ville où ils se déroulent.

Il est essentiel de connaître exactement la forme du géoïde pour dresser des cartes exactes, et jusqu'aux années cinquante cela n'a été fait que pour 7 % de la surface du globe. A ce moment-là, par exemple, la distance entre Londres et New York n'était connue qu'à un ou deux kilomètres près, et pour certaines îles du Pacifique cette incertitude pouvait atteindre dix kilomètres. En notre époque de voyages aériens et (hélas !) de pointage de missiles, ces incertitudes sont gênantes. Depuis, on a pu réaliser des cartes vraiment exactes – curieusement par des mesures astronomiques d'un nouveau type, et non par arpentage précis à la surface de la Terre. Le premier instrument de ces nouvelles mesures a été le satellite artificiel *Vanguard I,* lancé par les États-Unis le 17 mars 1958. *Vanguard I* tournait autour de la Terre en deux heures et demie, et a fait ainsi autant de tours en un ou deux ans que la Lune en a fait

depuis l'invention de la lunette. En observant la position du satellite à des instants précis, depuis des points précis de la Terre, on a pu calculer les distances entre ces points. De cette manière, la précision de ces mesures de distances est passée en quelques années de l'ordre du kilomètre à celui de la centaine de mètres. (Un autre satellite, *Transit I-B*, lancé par les États-Unis le 13 avril 1960, a inauguré une série spécialement destinée à étendre la méthode, pour préciser les positions à la surface de la Terre, aidant ainsi considérablement la navigation maritime et aérienne.)

Comme la Lune, *Vanguard I* décrit une ellipse située en dehors du plan équatorial de la Terre. Et comme dans le cas de la Lune, son périgée (point le plus proche du centre de la Terre) se déplace à cause de l'attraction du renflement équatorial. Mais *Vanguard I* est beaucoup plus proche de ce renflement, et beaucoup plus petit que la Lune ; il en est plus affecté. De plus, la cadence rapide de ses révolutions permet une étude plus facile. Dès 1959, on s'est rendu compte que la dérive du périgée de *Vanguard I* n'était pas la même dans l'hémisphère nord et dans l'hémisphère sud, et donc que le renflement n'était pas symétrique par rapport à l'équateur : il semblait plus haut de huit mètres (c'est-à-dire plus éloigné du centre de la Terre de la même distance) en des points situés au sud de l'équateur qu'en des points situés au nord. D'autres calculs ont montré que le pôle Sud était quinze mètres plus près du centre de la Terre que le pôle Nord (le tout ramené au niveau de la mer, bien entendu).

Des résultats ultérieurs, obtenus en 1961 à partir des orbites de *Vanguard I* et *Vanguard II* (lancé le 17 février 1959), montrent que l'équateur n'est pas un cercle parfait. A certains endroits, le diamètre équatorial a près de 400 mètres de plus qu'à d'autres endroits.

Les journaux ont parlé à ce propos de « Terre en forme de poire » et d'« équateur ovoïde ». En réalité, ces écarts par rapport à la perfection ne sont perceptibles qu'avec des moyens extrêmement précis. En regardant la Terre depuis l'espace, on ne verrait ni une poire ni un œuf : on verrait une sphère parfaite. De plus, les études précises du géoïde ont révélé tant de régions de léger aplatissement ou au contraire de léger renflement que s'il fallait décrire en un mot la forme de la Terre, il faudrait la qualifier de « bosselée »...

Finalement les satellites, y compris par des méthodes aussi directes que la photographie de la surface, ont permis de dresser la carte du monde entier avec une précision de l'ordre du mètre.

Les avions et les bateaux, qui déterminaient leur position en se repérant par rapport aux étoiles, le font maintenant grâce aux signaux émis par les *satellites de navigation*, sans avoir besoin d'un ciel clair, puisque les ondes radio traversent nuages et brouillard. Même des sous-marins en plongée peuvent en faire autant. Et la précision est telle que sur un paquebot on peut mesurer ainsi la distance entre la passerelle et la cuisine !

PESER LA TERRE

Connaissant exactement la forme et les dimensions de la Terre, on peut calculer son volume, qui vaut environ 1 080 milliards de kilomètres cubes. Pour calculer sa masse, c'est plus difficile, mais la loi de gravitation de

Newton nous donne un point de départ. Suivant cette loi, la *force d'attraction gravitationnelle F* entre deux objets quelconques de masses m_1 et m_2, séparés par une distance d, vaut :

$$F = \frac{Gm_1m_2}{d^2}$$

G étant *la constante universelle de gravitation.*

Newton ne pouvait pas donner la valeur de cette constante. Mais si l'on parvient à mesurer les autres termes de l'équation, on peut la calculer :

$$G = \frac{Fd^2}{m_1m_2}$$

Pour cela, il faut pouvoir mesurer la force d'attraction entre deux objets de masses connues séparés par une distance connue. Malheureusement, la force de gravitation est très faible, et entre deux objets de masses raisonnables, que nous puissions manipuler, elle est presque impossible à mesurer.

Cependant, en 1798, le physicien anglais Henry Cavendish réussit à y parvenir. C'était un riche excentrique, qui vivait dans un isolement presque complet, et réalisa quelques-unes des plus fines expériences de l'histoire des sciences. Pour mesurer G, il fixa deux boules de masse connue aux extrémités d'une longue tige, et suspendit le tout à un fil très fin. Puis il approcha de chaque boule une boule plus grosse, de masse également connue, dans le plan horizontal de la tige et de part et d'autre de celle-ci (figure 4.1), de manière que les deux attractions (de la grosse boule sur la petite) s'ajoutent pour faire tourner la tige et tordre le fil. Le système tournait effectivement d'un petit angle. Cavendish le mesura, puis détermina la force nécessaire pour tordre le fil du même angle. Il en déduisit F. Connaissant m_1 (masse de la grosse boule), m_2 (masse de la petite) et d (distance entre les deux), il calcula G. Dès lors, pour connaître la masse de la Terre, il lui suffisait de mesurer l'attraction qu'elle exerce sur un objet – c'est-à-dire le poids de cet objet : Cavendish pesa la Terre !

Au cours du XIXe siècle, on améliora la précision de la détermination de G. La valeur moderne, calculée en 1942 par P.R. Heyl et P. Chizanowski au Bureau national américain des poids et mesures, vaut $6{,}673\ 10^{-11}$ MKS. La petitesse de cette valeur est telle que deux masses d'un kilo, placées à un mètre l'une de l'autre, s'attirent avec une force égale au poids d'une particule de quelques milliardièmes de gramme.

Or, l'une de ces masses d'un kilo est attirée par la Terre avec une force qui est précisément le poids d'un kilo, ce qui donne une idée de la masse de la Terre – sans oublier que cette fois la distance (entre l'objet et le centre de la Terre) vaut 6 400 kilomètres. En fait, la masse de la Terre est 6 000 000 000 000 000 000 000 000 kilos, ce que l'on peut écrire plus simplement $6\ 10^{24}$ kilos.

Connaissant la masse de la Terre et son volume, on calcule facilement sa densité : 5,5 fois celle de l'eau. Or, la densité des roches de la surface est de l'ordre de 2,8, ce qui montre que celle de l'intérieur doit être beaucoup plus grande. Augmente-t-elle régulièrement vers le centre ? Non,

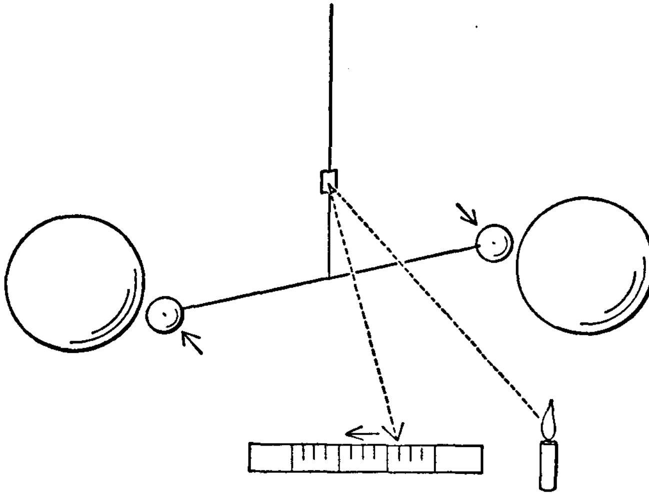


Figure 4.1. L'appareil de Cavendish pour mesurer la gravitation. Les deux grosses boules attirent les deux petites, tordant le fil de suspension. Le miroir permet de mesurer le petit angle de torsion, grâce au déplacement sur la graduation du faisceau lumineux réfléchi.

et la première preuve qu'elle ne le fait pas, et que la Terre est formée d'une série de couches distinctes, est fournie par l'étude des tremblements de terre.

La structure intérieure

LES TREMBLEMENTS DE TERRE

Il n'y a pas beaucoup de catastrophes naturelles capables de tuer des centaines de milliers de personnes en cinq minutes : presque uniquement les tremblements de terre.

Il y en a en moyenne un million par an, dont au moins cent sérieux et une dizaine de catastrophiques. Le pire de tous a probablement eu lieu dans le nord de la Chine en 1556, causant la mort de 830 000 personnes. Ensuite, le 30 décembre 1703, 200 000 personnes ont péri à Tokyo, et le 11 octobre 1737, 300 000 à Calcutta.

A cette époque-là, la science se développait en Europe occidentale, où on ne se préoccupait pas beaucoup de ce qui se passait de l'autre côté du monde. C'est alors qu'un désastre allait se produire en Europe même.

Le 1er novembre 1755, un tremblement de terre très violent, peut-être le plus violent des temps modernes, frappa la ville de Lisbonne, détruisant complètement la ville basse. Puis une vague géante – ce qu'on appelle un *tsunami* – arriva de l'océan. Enfin, deux nouveaux chocs se produisirent,

et de nombreux incendies éclatèrent. Il y eut 60 000 victimes, et la ville fut complètement dévastée.

On ressentit le choc sur des millions de kilomètres carrés, et les destructions furent importantes dans tout le Portugal et jusqu'au Maroc. Comme c'était la Toussaint, les églises étaient pleines, et dans toute l'Europe du Sud les fidèles virent danser les lustres.

Le désastre de Lisbonne fit une grande impression sur tout le monde savant. A cette époque optimiste, la jeune science de Galilée et de Newton semblait promettre aux hommes les moyens de faire de la Terre un paradis. Et voilà qu'était cruellement rappelée l'existence de forces colossales, imprévisibles et néfastes, qui échappaient au contrôle des hommes. C'est ce tremblement de terre qui inspira à Voltaire sa célèbre satire pessimiste *Candide*, avec son refrain ironique : tout est pour le mieux dans le meilleur des mondes...

Quand nous évoquons un tremblement de terre, nous pensons à ce qui se produit sur la terre ferme. Mais le fond des mers tremble aussi, et avec des effets encore plus désastreux. L'ébranlement déclenche en effet une très longue houle, et quand elle arrive près des côtes – surtout dans les étranglements formés par les ports – cette houle déferle en vagues monstrueuses, atteignant des dizaines de mètres de haut. Si rien ne permet de prévoir leur arrivée, des milliers de personnes sont noyées. Comme les côtes japonaises sont particulièrement éprouvées par ces vagues, on leur donne un nom japonais : *tsunamis* (qui veut dire « vagues dans le port »).

Après le désastre de Lisbonne, auquel un tsunami avait largement contribué, le monde scientifique se pencha sérieusement sur les causes possibles des tremblements de terre. Or, sur ce point, la meilleure réponse trouvée par les anciens Grecs (après les idées archaïques qui y voyaient les soubresauts des géants emprisonnés sous terre) est citée par Aristote : des masses d'air, enfermées sous terre, tenteraient de s'échapper. Pourtant cette explication ne satisfaisait pas les scientifiques des temps modernes, qui soupçonnaient plutôt des tensions dans les roches, provoquées peut-être par la chaleur interne de la Terre.

Le géologue anglais John Michell (qui avait étudié les forces de torsion, plus tard utilisées par Cavendish pour « peser la Terre ») suggéra en 1760 que les tremblements de terre étaient des ondes déclenchées par les mouvements de masses de roches à des kilomètres de profondeur. Il fut aussi le premier à penser que les tsunamis résultaient de tremblements de terre sous-marins.

Pour étudier les tremblements de terre, il fallait donc un instrument capable de détecter et de mesurer ces ondes, et il ne fut mis au point que cent ans après le désastre de Lisbonne : c'était le *sismographe* (du grec signifiant « tremblement de terre » et « écrire »), réalisé en 1855 par le physicien italien Luigi Palmieri.

L'appareil de Palmieri était formé d'un tube horizontal, coudé à la verticale aux deux bouts, et contenant du mercure. Quand le sol tremblait, le mercure se déplaçait. L'appareil réagissait bien aux tremblements de terre, mais malheureusement aussi à d'autres ébranlements – celui d'un chariot dans la rue, par exemple.

En 1880, un ingénieur anglais, John Milne, réalisa un appareil beaucoup plus perfectionné, qui est l'ancêtre de tous les sismographes actuels. Enseignant la géologie à Tokyo, Milne en a profité pour étudier les

tremblements de terre, très fréquents au Japon, et son appareil est le fruit de ces études.

Sous la forme la plus simple, le sismographe de Milne se compose d'un bloc massif attaché par un ressort relativement faible à un support rigidement lié au rocher. Quand la Terre bouge, le bloc reste immobile, à cause de son inertie, et le ressort se contracte ou s'étire un peu, suivant les mouvements du rocher. Ce déplacement relatif est enregistré sur un tambour qui tourne lentement, au moyen d'une plume fixée au bloc, et qui écrit sur un papier couvert de noir de fumée. En réalité, on utilise deux blocs, l'un orienté de manière à enregistrer les ondes nord-sud et l'autre les ondes est-ouest. Les vibrations « parasites », dont l'origine n'est pas dans la masse rocheuse, n'affectent pas le sismographe. Actuellement, les sismographes les plus sensibles, comme celui de l'université de Fordham, utilisent un rayon lumineux au lieu d'une plume, pour éviter le frottement de celle-ci sur le papier. Le rayon impressionne un papier sensible à la lumière.

Milne a joué un rôle important dans l'établissement de stations destinées à étudier les tremblements de terre et les phénomènes qui y sont liés, dans diverses parties du monde et principalement au Japon. Dès 1900, treize stations sismographiques fonctionnaient, et maintenant il y en a plus de cinq cents, réparties sur tous les continents, y compris l'Antarctique. Moins de dix ans après l'établissement des premières stations, la théorie de Michell, suivant laquelle les tremblements de terre étaient causés par des ondes qui se propagent à travers la masse du globe, était amplement vérifiée.

Si on comprend mieux les tremblements de terre, ils ne se produisent pas moins souvent pour autant, et ne sont pas moins désastreux quand ils se produisent. Les années soixante-dix, par exemple, ont connu plusieurs tremblements de terre très sérieux.

Le 27 juillet 1976, un tremblement en Chine a détruit une ville au sud de Pékin, faisant environ 650 000 morts, le plus meurtrier de tous les tremblements recensés depuis celui de Chine du Nord quatre siècles plus tôt. Pendant la même période, des tremblements de terre désastreux se sont produits aussi au Guatemala, au Mexique, en Italie, aux Philippines, en Roumanie et en Turquie.

Ces nombreuses catastrophes ne signifient pas que la Terre devient moins stable qu'avant. Les moyens de communication modernes nous signalent immédiatement les tremblements de terre – souvent avec des reportages télévisés en direct – alors qu'il y a quelques dizaines d'années seulement, des cataclysmes lointains seraient passés inaperçus. De plus, les tremblements de terre ont plus de chances d'être catastrophiques qu'il y a seulement cent ans : la population du globe a considérablement augmenté, elle s'entasse dans des villes immenses, et les constructions humaines – vulnérables aux tremblements de terre – sont devenues beaucoup plus nombreuses.

Ce sont là d'autant plus de raisons pour mettre au point des méthodes permettant de prévoir les tremblements de terre. Les sismologues cherchent pour cela à repérer des indices significatifs : le sol pourrait s'élever, les roches pourraient s'écarter ou se rapprocher, absorbant ou exprimant de l'eau, ce qui ferait varier le niveau dans les puits. Le magnétisme ou la conductivité électrique des roches pourrait changer. Les animaux, sensibles à des vibrations très faibles, ou à des changements de

l'environnement imperceptibles pour nous, pourraient commencer à réagir en donnant des signes de nervosité.

Les Chinois, en particulier, se sont mis à recueillir tous les rapports d'événements inhabituels, même l'écaillement des peintures, et selon leurs rapports, ont pu prédire un tremblement de terre dans le Nord-Est de la Chine le 4 février 1975. La population a pu quitter la ville à temps et se réfugier en rase campagne, ce qui a évité des milliers de victimes. Mais le tremblement de terre beaucoup plus violent et désastreux de 1976, lui, n'a pas du tout été prévu.

Ces prédictions posent un autre problème, tant qu'elles ne sont pas encore bien sûres : elles peuvent faire plus de mal que de bien. Une fausse alerte peut désorganiser la vie économique, et provoquer plus de dégâts que n'en aurait fait un tremblement de terre bénin. De plus, après une ou deux fausses alertes, un avertissement sérieux pourrait fort bien être ignoré.

Les dégâts causés par les tremblements de terre ne sont pas surprenants : les plus violents libèrent une énergie de l'ordre de celle de cent mille bombes A, ou de cent grosses bombes H. Cette énergie se répartit sur de très grandes surfaces, sans quoi les tremblements de terre seraient encore plus meurtriers. Ils font entrer en vibration la Terre entière, comme un gigantesque diapason. Le tremblement de terre du Chili, en 1960, a fait vibrer la planète avec une période tout juste inférieure à une heure, période beaucoup trop longue, évidemment, pour correspondre à un son audible.

On mesure l'intensité des tremblements de terre sur une échelle de 0 à 9, chaque numéro correspondant à une énergie libérée environ 31 fois plus grande que le précédent. On n'a jamais enregistré de tremblements de terre d'intensité supérieure à 9, mais celui du Vendredi saint 1964, en Alaska, a atteint une intensité de 8,5. Cette échelle s'appelle *l'échelle de Richter,* du nom du sismologue américain qui l'a proposée en 1935.

Heureusement, toutes les régions de la Terre ne sont pas également menacées (encore que cela ne réconforte pas ceux qui habitent précisément dans les régions à haut risque...).

80 % de l'énergie libérée par les tremblements de terre l'est sur les rives de l'océan Pacifique, et 15 % suivant une bande est-ouest qui traverse la Méditerranée. Ces zones (voir figure 4.2) coïncident étroitement avec des régions volcaniques – d'où l'idée déjà ancienne que les tremblements de terre sont liés à la chaleur interne.

LES VOLCANS

Les éruptions volcaniques sont des phénomènes naturels aussi effrayants que les tremblements de terre, et de durée plus longue, mais la plupart du temps leurs effets sont limités à des régions très restreintes. On a répertorié dans l'histoire environ cinq cents volcans actifs à un moment ou à un autre, dont les deux tiers sont au bord du Pacifique.

Exceptionnellement, un volcan peut emprisonner et surchauffer une grande quantité d'eau, et alors la catastrophe est épouvantable. Le 26 et le 27 août 1883, la petite île de Krakatoa, située dans le détroit qui sépare Java de Sumatra, a explosé avec un bruit qui fut sans doute le plus fort jamais produit sur Terre, au moins depuis le début des temps historiques. On l'a entendu à cinq mille kilomètres et les instruments ont pu le détecter

d'un bout à l'autre du monde. Les ondes sonores ont fait plusieurs fois le tour du monde. Vingt kilomètres cubes de roches ont été réduits en pièces et projetés dans les airs, retombant sur une surface de près d'un million de kilomètres carrés. Les cendres ont obscurci le ciel sur des centaines de kilomètres carrés, laissant dans la stratosphère des traces qui ont donné pendant deux ans des couchers de Soleil exceptionnellement colorés. Des tsunamis de trente mètres de haut ont fait 36 000 victimes sur les côtes de Java et de Sumatra, donnant des vagues facilement détectables dans le monde entier.

Un événement analogue, avec des conséquences encore plus considérables, s'était peut-être produit trois mille ans plus tôt dans la Méditerranée. En 1967, des archéologues américains ont découvert des ruines ensevelies sous les cendres dans la petite île de Santorin, à cent trente kilomètres au nord de la Crète. Il semble que, environ en 1400 av. J.-C., cette île ait explosé comme Krakatoa, mais avec plus de force encore, un bruit peut-être encore plus grand, et des conséquences encore plus désastreuses. Le tsunami a en effet frappé la Crète, qui était alors le siège d'une civilisation parmi les plus avancées de ce temps, coup terrible dont cette civilisation ne devait jamais se remettre : la Crète perdit le contrôle des mers, et devait connaître plusieurs siècles de désordre et d'obscurité avant de retrouver une partie de sa grandeur passée. Cette catastrophe, dont le récit à été embelli de génération en génération, est peut-être à l'origine de l'histoire de l'Atlantide, racontée par Platon plus de mille ans après la mort de Santorin et de la civilisation crétoise.

Mais l'éruption la plus célèbre de l'histoire du monde est insignifiante, comparée à celle de Krakatoa ou à celle de Santorin : c'est l'éruption du Vésuve en l'an 79 av. J.-C.

Jusque-là, le Vésuve avait été considéré comme éteint. Son réveil ensevelit les villes romaines de Pompéi et d'Herculanum. Le célèbre encyclopédiste Pline l'Ancien trouva la mort dans cette catastrophe, décrite par son neveu, Pline le Jeune, l'un des survivants.

Les fouilles de ces deux sites ont commencé de façon sérieuse après 1763. Elles ont permis d'étudier, dans des conditions exceptionnelles de préservation, le cadre de vie au moment le plus florissant de l'Antiquité romaine.

Autre phénomène exceptionnel : la naissance d'un nouveau volcan. Le 20 février 1943, dans le village de Paricutin, à l'ouest de Mexico, un volcan est apparu dans ce qui était jusque-là un champ de maïs comme les autres. En huit mois, le cône de cendres a atteint cinq cent mètres de haut. Bien sûr, le village a dû être abandonné.

Dans l'ensemble, les Américains ne se préoccupent pas beaucoup des éruptions volcaniques : elles se produisent en général à l'étranger. Néanmoins, le plus grand volcan actif du monde se trouve à Hawaï, qui fait partie des États-Unis. C'est le Kilauea, dont le cratère mesure dix kilomètres carrés, et qui est fréquemment en éruption. Mais ses éruptions ne sont pas explosives : la lave déborde, mais elle se déplace assez lentement pour qu'il n'y ait pas de victimes, même si elle provoque parfois des dégâts matériels. Ce volcan a été particulièrement actif en 1983.

La chaîne des Cascades, qui suit la côte pacifique des États-Unis, à deux cents kilomètres vers l'intérieur, depuis la Californie du Nord jusqu'au Canada, possède de nombreux volcans éteints, dont certains sont très célèbres, comme le mont Rainier et le mont Hood. Etant éteints, ils

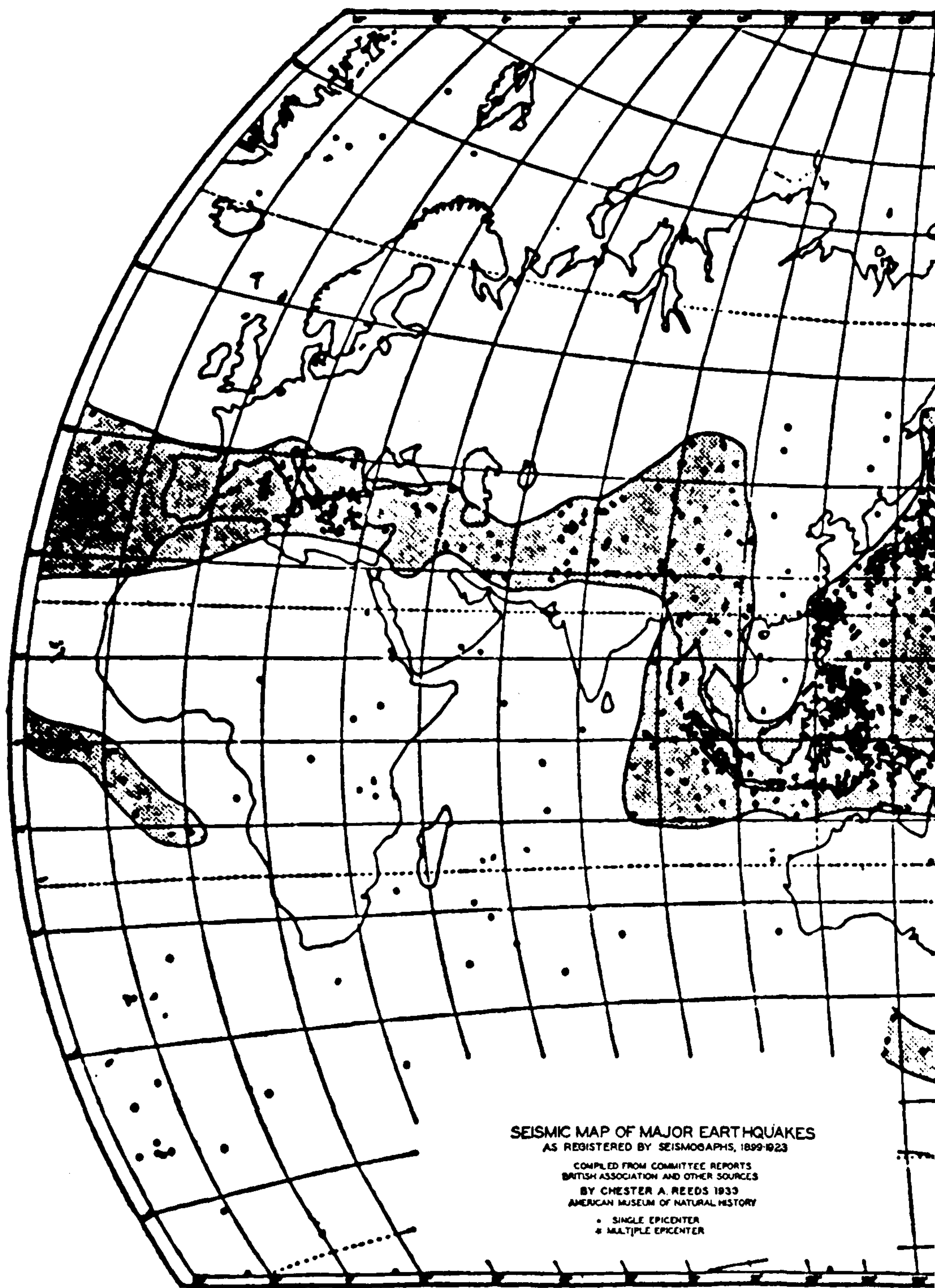


Figure 4.2. Les épicentres des tremblements de terre entre 1899 et 1923. (Avec l'aimable autorisation du Centre géographique national, Administration nationale de l'Océan et de l'atmosphère, États-Unis.)



n'inspirent aucune inquiétude ; pourtant un volcan peut dormir pendant des siècles, et se réveiller tout à coup.

Les Américains s'en sont rendu compte à propos du mont Saint-Helens, dans le nord de la même chaîne. Il avait été actif entre 1831 et 1854, mais à l'époque la région était peu peuplée, et nous n'avons pas beaucoup de détails sur cette activité. Ensuite, le volcan était resté tranquille pendant plus d'un siècle, avant de se réveiller brusquement le 18 mai 1980. Ce jour-là, après quelques grondements et quelques secousses, il expulsa soudain une grande quantité de cendres, faisant une centaine de victimes, pour une bonne part des imprudents qui avaient refusé de s'éloigner. Ce ne fut pas une éruption catastrophique, mais elle a eu un certain retentissement, parce que c'était la première aux États-Unis (tout au moins dans leur partie continentale, et en excluant l'Alaska) depuis de nombreuses années.

Le danger des éruptions n'est pas seulement immédiat. Celles qui sont très importantes projettent dans la haute atmosphère une grande quantité de poussières, qui peuvent mettre des années à retomber. Après l'éruption du Krakatoa, les couchers de Soleil ont été exceptionnels dans le monde entier pendant deux ans, à cause de la diffusion de la lumière par les poussières en suspension. Mais, chose moins innocente, la poussière peut diminuer sérieusement la quantité de soleil reçue par la surface terrestre.

Il arrive que cet effet soit local, mais catastrophique. En 1783, le volcan Laki, en Islande, entra en éruption. En deux ans, la lave couvrit presque mille kilomètres carrés, mais sans causer de grands dommages. Toutefois, l'éruption projeta de la poussière et du gaz sulfureux sur presque toute l'Islande (et jusqu'en Ecosse). La poussière obscurcit le ciel, ce qui fit périr les récoltes. Le gaz sulfureux tua les trois quarts du bétail de toute l'île. Aussi dix mille personnes, un cinquième de la population, moururent de faim et de maladie.

Le 7 avril 1815, le mont Tambora, situé sur une petite île à l'est de Java, projeta dans la haute atmosphère cent cinquante kilomètres cubes de pierres et de poussières. Les températures sur toute la Terre furent plus basses que d'habitude pendant un an. En Nouvelle-Angleterre, il gela tous les mois, même en juillet et en août : on parle de l'année 1816 comme de « l'année sans été ».

Il arrive aussi que les volcans tuent immédiatement, sans lave et même sans cendres. Le 8 mai 1902, la montagne Pelée, à la Martinique, cracha soudain un épais nuage de gaz brûlant, qui descendit rapidement les pentes, droit vers la ville de Saint-Pierre, la plus importante de l'île. En trois minutes, les 38 000 habitants étaient morts, asphyxiés. Le seul survivant fut un condamné à mort, qui attendait l'exécution dans un cachot souterrain...

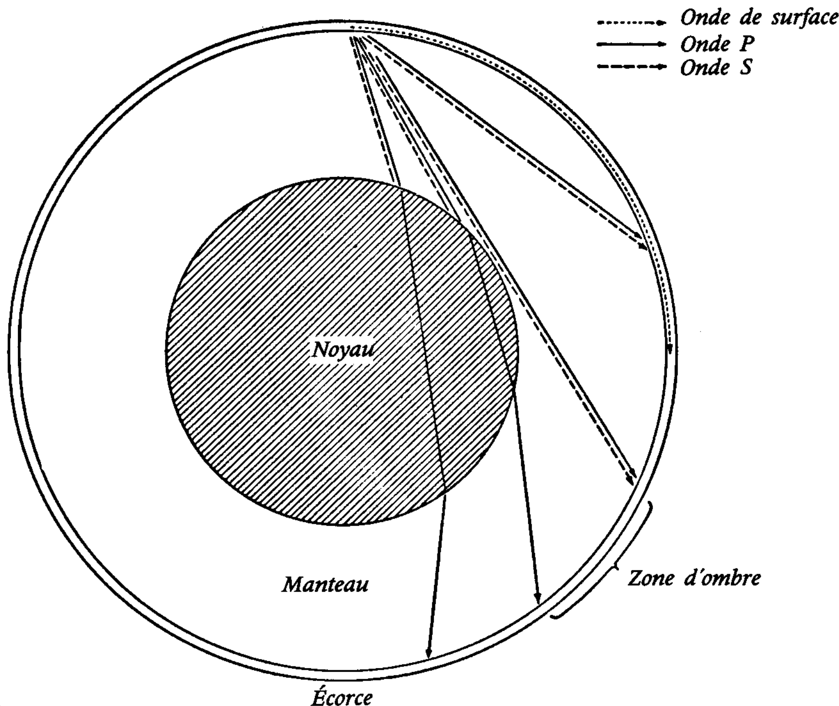
FORMATION DE LA CROÛTE TERRESTRE

La recherche moderne sur les volcans et leur rôle dans la formation d'une grande partie de la croûte terrestre a commencé au milieu du XVIIIe siècle, avec le géologue français Jean-Etienne Guettard. Ensuite, pendant les dernières années du siècle, les efforts isolés du géologue allemand Abraham Gottlob Werner ont popularisé l'idée fausse suivant laquelle la plupart des roches étaient d'origine sédimentaire (*neptunisme*).

Mais les résultats concrets présentés, en particulier, par Hutton, ont montré sans aucun doute que la plupart des roches avaient une origine volcanique (*plutonisme*). Volcans et tremblements de terre semblent manifester l'énergie interne de la Terre, résultant surtout de la radioactivité (voir le chapitre 7).

Une fois que les sismographes ont permis l'étude détaillée des ondes liées aux tremblements de terre, on a constaté que les plus faciles à étudier se rangeaient en deux catégories : les *superficielles* et les *profondes*. Les premières suivent la courbure de la Terre, les autres traversent l'intérieur – raccourci qui leur permet en général d'arriver les premières au sismographe. Ces ondes profondes se répartissent à nouveau en deux types : les ondes primaires (*ondes P*) et secondaires (*ondes S*). Les premières sont longitudinales, commes les ondes sonores, et se traduisent par des compressions et détentes alternées du milieu qu'elles traversent. Elles peuvent voyager dans n'importe quel milieu, solide ou fluide. Les ondes secondaires, au contraire, sont transversales : la direction des oscillations est à angle droit avec celle de la propagation. Ces ondes ne peuvent pas traverser les liquides ou les gaz. Les ondes primaires se déplacent plus vite que les secondaires. Elles arrivent donc un peu avant au sismographe. Ce décalage permet d'estimer la distance à laquelle s'est produite la secousse. Si on mesure cette distance à partir de plusieurs stations, et si on trace

Figure 4.3. Parcours des ondes engendrées par un tremblement de terre. Les ondes de surface voyagent le long de l'écorce. Le noyau liquide de la Terre réfracte les ondes P. Les ondes S ne peuvent pas traverser le noyau.



à chaque fois un cercle centré sur la station et ayant pour rayon la distance estimée, tous les cercles vont se couper en un point, appelé *épicentre* du tremblement de terre, et situé à la surface, à la verticale du point où s'est produite la secousse.

Les vitesses des deux types d'ondes dépendent du type de roche traversée, de la température et de la pression, comme l'ont montré des études en laboratoire. On peut donc se servir de ces ondes comme « sondes » pour étudier les conditions loin au-dessous de la surface.

Près de la surface, une onde primaire se déplace à 8 km/s. 1 500 kilomètres plus bas, d'après les temps d'arrivée, sa vitesse doit être proche de 13 km/s. De la même façon, une onde secondaire va à moins de 5 km/s près de la surface, et à 7 km/s à 1 500 kilomètres de profondeur. Or, l'augmentation de vitesse correspond à une augmentation de densité, et on peut ainsi estimer la densité de la roche en profondeur. A la surface de la Terre, nous l'avons vu, la densité moyenne est 2,8. 1 500 kilomètres plus bas, elle est de 5, et à 2 900 kilomètres au-dessous de la surface, la densité atteint 6.

A cette profondeur de 2 900 kilomètres, un changement brutal se produit. Les ondes secondaires ne passent plus. Le géologue anglais Richard Dixon Oldham en a déduit, en 1906, que la région située en dessous de cette limite devait être liquide. En atteignant la limite de ce *noyau liquide* de la Terre, les ondes secondaires sont arrêtées, et les primaires changent brusquement de direction : elles sont apparemment réfractées en entrant dans le noyau liquide.

Cette limite s'appelle *discontinuité de Gutenberg*, du nom du géologue américain Beno Gutenberg, qui a défini cette limite en 1914, et a montré que le noyau avait un rayon de 3 500 kilomètres. En 1936, le mathématicien australien Keith Edward Bullen a calculé, à partir des résultats sismographiques disponibles, les densités des couches profondes de la Terre. Ses résultats ont été confirmés lors du grand tremblement de terre du Chili en 1960. On peut donc dire qu'à la discontinuité de Gutenberg, la densité passe brusquement de 6 à 9, et augmente ensuite régulièrement pour atteindre au centre la valeur 11,5.

LE NOYAU LIQUIDE

Quelle est la nature du noyau ? Il doit être fait d'un matériau dont la densité va de 9 à 11,5 dans les conditions de température et de pression aux différents points du noyau. On estime à 1 360 000 atmosphères la pression à la surface du noyau, et à 3 400 000 atmosphères la pression au centre. La température est plus difficile à estimer. A partir de la cadence à laquelle elle augmente avec la profondeur dans les mines, et de la conductivité thermique des roches, on suppose (en gros) que la température du noyau doit être au moins de 5 000 °C (le centre d'une planète beaucoup plus grosse, comme Jupiter, atteint peut-être des températures de l'ordre de 50 000 °C).

Le noyau doit être formé d'un élément assez répandu pour former une sphère ayant la moitié du diamètre de la Terre et le tiers de sa masse. Or, le seul élément lourd qui soit commun dans l'Univers est le fer. A la surface de la Terre, sa densité n'est que de 7,86, mais sous les pressions énormes qui règnent dans le noyau, cette densité doit avoir des valeurs convenables (entre 9 et 12). De plus, dans ces conditions, le fer est liquide.

Un argument supplémentaire est fourni par les météorites. Celles-ci sont de deux sortes : les *météorites rocheuses*, formées surtout de silicates, et les *météorites de fer*, qui contiennent 90 % de fer, 9 % de nickel, et 1 % d'autres éléments. On pense généralement que les météorites sont les débris provenant de l'éclatement d'astéroïdes, dont certains auraient été assez gros pour que leurs parties rocheuses et métalliques soient distinctes. Dans ce cas, les parties métalliques sont évidemment composées de fer au nickel, ce qui incite à penser que le noyau terrestre l'est aussi. D'ailleurs dès 1866, bien avant l'exploration profonde par les sismologues, la composition des météorites avait suggéré au géologue français Gabriel Auguste Daubrée que le noyau de la Terre pourrait bien être formé de fer.

Actuellement, la plupart des géologues considèrent le noyau liquide de fer au nickel comme allant de soi. Mais on y a apporté un perfectionnement important. En 1936 le géologue danois Inge Lehmann, cherchant à expliquer l'arrivée de quelques ondes primaires dans la « zone d'ombre » de la surface – en principe inaccessible (voir figure 4.3) – a émis l'hypothèse d'une discontinuité dans le noyau, à 1 300 kilomètres du centre. Cette discontinuité pouvait dévier une faible partie des ondes dans la zone d'ombre. Cette idée a été soutenue par Gutenberg, et actuellement la plupart des géologues distinguent un *noyau externe* de fer au nickel sous forme liquide et un *noyau interne* qui en diffère d'une manière ou d'une autre, soit en étant solide, soit en ayant une composition chimique légèrement différente. Lors du grand tremblement de terre du Chili, en 1960, le globe tout entier est entré en vibration, à des fréquences très basses en accord avec les prédictions faites en tenant compte du noyau interne, ce qui constitue un argument puissant en faveur de son existence.

LE MANTEAU

On appelle ainsi tout ce qu'il y a autour du noyau de la Terre. Ce manteau semble formé de silicates, mais à en juger par la vitesse des ondes qui les traversent, ces silicates ne sont pas les mêmes que ceux qui forment les roches de la surface – remarque signalée pour la première fois en 1919 par le physico-chimiste américain Leason Heberling Adams. Leurs propriétés laissent penser qu'ils contiennent beaucoup de magnésium et de fer, et peu d'aluminium.

Le manteau proprement dit ne s'étend pas tout à fait jusqu'à la surface. Le géologue croate Andrija Mohorovicic, étudiant les ondes produites par un tremblement de terre dans les Balkans en 1909, découvrit l'existence d'une brusque augmentation de leur vitesse à une trentaine de kilomètres sous la surface. On appelle maintenant cette limite de la *croûte* terrestre *discontinuité de Mohorovicic* ou par abréviation *Moho*.

C'est grâce aux ondes de surface, que nous avons mentionnées plus haut, que l'on peut étudier le plus facilement la croûte et le manteau supérieur. Comme les ondes profondes, les ondes de surface se divisent en deux catégories : les *ondes de Love* (du nom de celui qui les a découvertes, Augustus Edward Hough Love) zigzaguent dans un plan horizontal, comme un serpent sur le sol ; les *ondes de Rayleigh* (d'après le physicien anglais John William Strutt, Lord Rayleigh) ondulent dans un plan vertical, comme le « serpent de mer » des légendes.

L'analyse de ces ondes de surface (en particulier par Maurice Ewing, de l'université Columbia) montre que la croûte a une épaisseur variable.

Elle est plus mince sous les océans, où le Moho remonte par endroits à une quinzaine de kilomètres seulement du niveau de la mer. Les océans atteignant 8 à 10 kilomètres de profondeur, il ne reste plus que quelques kilomètres d'épaisseur pour la croûte. Sous les continents, au contraire, le Moho se trouve à une trentaine de kilomètres sous le niveau de la mer (35 kilomètres au-dessous de New York, par exemple), et descend jusqu'à 60 kilomètres sous les chaînes de montagnes. Cela, s'ajoutant aux résultats des mesures du champ de gravitation à la surface, permet d'estimer que les roches formant les chaînes de montagnes sont plus légères que la moyenne.

On arrive ainsi à une image de la croûte où apparaissent deux matériaux principaux, le basalte et le granit, le second – moins dense – flottant sur le premier pour former les continents, et, là où l'épaisseur de granit est particulièrement importante, des montagnes (de même qu'un gros iceberg s'élève plus haut au-dessus de l'eau qu'un petit). Les jeunes montagnes enfoncent leurs « racines » de granit profondément dans le basalte, mais à mesure que l'érosion les use, elles maintiennent leur équilibre isostatique (terme proposé en 1889 par le géologue américain Clarence Edward Dutton) en flottant un peu plus haut. Sous les très vieilles chaînes de montagnes, comme les Appalaches, il n'y a presque plus de racines.

Sous les océans, le basalte est couvert d'une couche sédimentaire de quelques centaines de mètres, mais ne supporte presque pas de granit : l'océan Pacifique n'en possède pas du tout. La minceur de la croûte sous les océans a suggéré un projet spectaculaire : pourquoi ne pas percer un trou à travers, pour voir de quoi est fait le manteau ? En réalité, ce serait difficile : il faudrait ancrer un navire précisément au-dessus d'une fosse abyssale, descendre le matériel de forage à travers dix kilomètres d'eau, et traverser ensuite une épaisseur de rocher plus grande que toutes celles que l'on aurait déjà percées. Aussi l'enthousiasme soulevé par ce projet est-il rapidement retombé, et on n'en parle plus beaucoup.

La « flottaison » du granit sur le basalte suggère inévitablement la possibilité de la *dérive des continents.* En 1912, le géologue allemand Alfred Lothar Wegener a publié une théorie suivant laquelle tous les continents auraient été initialement rassemblés en un seul morceau de granit, qu'il a appelé Pangée (« terre globale » en grec). Assez tôt dans l'histoire de la Terre, ce radeau se serait fracturé, et les continents auraient dérivé de plus en plus loin les uns des autres. Cette dérive continuerait, le Groenland, par exemple, s'éloignant de l'Europe à une vitesse d'un mètre par an. L'un de ses nombreux arguments (déjà pressenti par d'autres, remontant jusqu'à Francis Bacon en 1620) était le fait que la côte orientale de l'Amérique du Sud « s'emboîte » exactement, comme une pièce de puzzle, dans la côte occidentale de l'Afrique.

Pendant cinquante ans, cette idée a été repoussée énergiquement par tout le monde. Moi-même, quand la première édition de ce livre a paru, en 1960, je me suis senti le droit, étant donné l'avis unanime des géophysiciens de l'époque, de la repousser catégoriquement. L'argument le plus fort contre la théorie de Wegener était la rigidité du basalte, trop grande pour permettre au granit de dériver en flottant dessus, même en disposant de millions d'années.

Mais peu à peu, les résultats se sont multipliés, rendant de plus en plus évident le fait que l'océan Atlantique n'avait pas toujours existé, et que les continents avaient commencé par former un seul bloc. Les sondages

ont révélé que les continents « s'emboîtaient » non seulement par leurs côtes, mais tout au long de la crête centrale de l'océan Atlantique, aussi bien au nord qu'au sud.

D'autre part, la structure des roches de l'Afrique occidentale s'accorde de manière détaillée avec celle des régions correspondantes – suivant Wegener – en Amérique du Sud. Enfin, les voyages des pôles magnétiques semblent moins étonnants si on considère que ce sont les continents qui se sont déplacés par rapport à eux, et non l'inverse.

A ces arguments géologiques en faveur de la Pangée et de sa rupture, sont venus s'ajouter des arguments biologiques encore plus évidents. En 1968, par exemple, on a trouvé dans l'Antarctique un os fossile (de six centimètres de long) ayant appartenu à un amphibien d'espèce disparue. Un tel animal n'aurait jamais pu vivre dans une région aussi froide, aussi près du pôle. Il n'aurait jamais pu traverser une étendue d'eau salée, si faible soit-elle. Il faut donc bien admettre que l'Antarctique a dû faire partie d'un bloc plus vaste, comportant des régions beaucoup plus chaudes. D'une manière générale, les renseignements fournis par les fossiles (dont nous parlerons au chapitre 16) corroborent fortement la théorie de Wegener, et l'existence, puis la rupture, de la Pangée.

Il est important de souligner ici la base de l'opposition rencontrée par Wegener de la part des autres géologues. Les gens qui s'agitent à la frange de la science tentent souvent de justifier leurs théories douteuses en accusant les scientifiques d'être dogmatiques et fermés aux idées nouvelles (ce qui peut être vrai dans certains cas, et à certains moments, mais jamais au point où ces « marginaux » le prétendent). Ils prennent souvent comme exemple l'attitude unanime d'opposition à la théorie de la dérive des continents de Wegener, et là ils ont tort.

Les géologues, en effet, ne se sont pas opposés à *l'idée* de la Pangée et de sa rupture. Ils ont même admis volontiers des *idées* beaucoup plus fantaisistes destinées à expliquer la diffusion de la vie à travers la surface de la Terre. Ce qui leur a servi à repousser en bloc la théorie de Wegener, c'était le mécanisme qu'il suggérait : la dérive de blocs de granit sur un océan de basalte. Sur ce point, leurs arguments se sont révélés justifiés : les continents ne dérivent pas *à travers* le basalte, mais avec le basalte, de même qu'un radeau ne dérive pas *à travers* l'eau, mais avec l'eau.

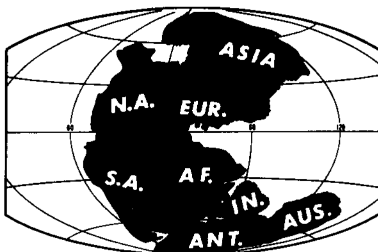
C'est ce mécanisme différent qui allait rendre compte des indications, géographiques et biologiques, des changements de position des continents – mécanisme plus plausible, et vérifié par l'expérience. Nous discuterons ces expériences dans la suite du chapitre ; disons seulement pour l'instant que vers 1960, le géologue américain Harry Hammond Hess a suggéré, sur la base de nouvelles découvertes, que la matière du manteau, fondue, pourrait jaillir – à travers des lignes de fracture s'étendant, par exemple, tout le long de l'Atlantique –, se refroidissant et se solidifiant à mesure, et provoquant ainsi un élargissement du plancher océanique. Les continents ne dérivent pas : ils sont entraînés peu à peu loin les uns des autres par l'élargissement des planchers océaniques.

Ainsi, on se rend compte maintenant que la Pangée a bien existé, somme toute, et existait encore il y a 225 millions d'années, au début du règne des dinosaures. A en juger par l'évolution et la distribution des plantes

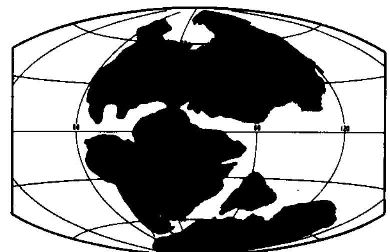
et des animaux, la rupture a dû se produire il y a environ 200 millions d'années, la Pangée se séparant alors en trois morceaux, ceux-là même qu'avait décrits Wegener : la Laurasie (Amérique du Nord, Europe et Asie), la partie méridionale (Amérique du Sud, Afrique et Inde), appelée Gondwana, du nom d'une province indienne, et une troisième partie, formée de l'Australie et de l'Antarctique.

Il y a 65 millions d'années, les dinosaures ayant fait place aux premiers mammifères, l'Amérique du Sud s'est séparée de l'Afrique, et l'Inde s'est détachée, pour se mettre en route vers l'Asie. Finalement, l'Amérique du Nord et l'Europe se sont séparées, et l'Inde est venue « emboutir » l'Asie (le choc faisant surgir l'Himalaya). L'Australie et l'Antarctique se sont séparés, et les continents se sont mis à ressembler à ce qu'ils sont actuellement (voir figure 4.4).

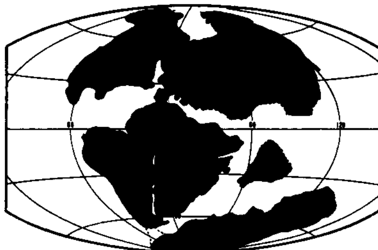
Figure 4.4. Les périodes géologiques.



Permien — 225 millions d'années



Trias — 200 millions d'années



Jurassique — 135 millions d'années



Crétacé — 65 millions d'années



Cénozoïque (époque actuelle)

L'ORIGINE DE LA LUNE

Une suggestion encore plus étonnante au sujet des changements qui ont pu prendre place sur la Terre au cours des époques géologiques est due à George Howard Darwin (fils de Charles Darwin). Cet astronome anglais a suggéré en 1879 que la Lune pourrait bien être un morceau de la Terre, qui se serait détaché il y a très longtemps, laissant un trou correspondant à l'océan Pacifique.

L'idée est séduisante, car la Lune forme à peu près un centième de la masse totale du couple Terre-Lune, et ses dimensions lui permettent de « tenir » dans le Pacifique. Si la Lune s'était vraiment formée à partir des couches externes de la Terre, cela expliquerait qu'elle soit beaucoup moins dense que la Terre (étant privée de noyau de fer), et que le plancher du Pacifique soit complètement dénué de granit.

Mais cette possibilité semble peu probable, pour différentes raisons, et actuellement il n'y a pratiquement aucun astronome ni aucun géologue pour l'admettre. Pourtant la Lune semble bien avoir été autrefois plus proche de la Terre qu'elle ne l'est aujourd'hui.

L'attraction gravitationnelle de la Lune produit des marées à la fois dans les océans et dans la croûte solide de la Terre. La Terre tournant sur elle-même, l'eau des océans se déplace sur les bas-fonds, et les couches de roche frottent les unes sur les autres en s'élevant et en s'abaissant. Ces frottements transforment peu à peu en chaleur une partie de l'énergie de rotation de la Terre, dont la période de rotation augmente. Cette augmentation de la durée du jour n'est pas très sensible pour nous : une seconde en 62 500 ans. Mais à mesure que la Terre ralentit son mouvement de rotation, la conservation du moment angulaire impose que la Lune augmente le sien : elle s'éloigne peu à peu de la Terre.

En remontant le temps vers le passé géologique le plus lointain, on trouve ainsi une rotation terrestre de plus en plus rapide, des jours de plus en plus courts, une Lune de plus en plus proche, ces modifications s'accélérant de plus en plus. George Darwin est remonté ainsi jusqu'au moment où, d'après lui, la Terre et la Lune ne formaient qu'un seul corps. Sans aller si loin, il semble bien que le jour devait être plus court dans le passé. Par exemple, il y a 570 millions d'années – l'âge des plus vieux fossiles – le jour devait durer 20 heures, et donc il devait y avoir 428 jours par an.

Or, ceci n'est plus de la pure théorie. On en a trouvé des preuves matérielles, en étudiant les coraux fossiles. Le corail travaille d'autant plus qu'il fait plus clair : plus le jour que la nuit, et plus l'été que l'hiver. Aussi, on y repère des renflements annuels, et de fines bandes qui correspondent aux jours successifs. En 1963, le paléontologiste américain John West Wells a compté ces bandes sur des coraux fossiles, et il a trouvé en moyenne 400 bandes fines (jours) par bande large (année) sur des coraux vieux de 400 millions d'années, et 380 jours par an sur des coraux vieux de 320 millions d'années.

Bien sûr, une question s'impose : si la Lune était plus proche de la Terre à cette époque, et si la Terre tournait plus vite, que se passait-il plus tôt encore ? Si la théorie de George Darwin est fausse, qu'est-ce qui est vrai ?

Plusieurs suggestions sont possibles. Suivant l'une d'elles, la Lune a été capturée par la Terre. Si cette capture a eu lieu il y a 600 millions d'années environ, cela expliquerait que l'on trouve beaucoup de fossiles à partir de cette époque, les roches plus anciennes ne présentant que de vagues

traces de carbone ; car elles ont peut-être été « lavées » par les marées gigantesques accompagnant la capture de la Lune (il n'y avait pas de vie terrestre à l'époque, sinon elle aurait été anéantie). Si cette hypothèse de capture de la Lune est correcte, et si elle était plus proche à cette époque, la récession de la Lune et l'allongement du jour ont commencé à ce moment-là et pas avant.

D'autres suggèrent que la Lune a pu se former au voisinage de la Terre, à partir du même nuage de poussières, et s'en éloigner depuis lors, mais sans jamais avoir été un morceau de la Terre.

On espérait que les échantillons lunaires rapportés par les astronautes dans les années soixante-dix permettraient de résoudre le problème, mais ils n'ont rien permis du tout. Par exemple, la surface de la Lune, contrairement à celle de la Terre, est couverte de morceaux de verre. La croûte lunaire est entièrement dépourvue d'eau, et pauvre en matériaux qui fondent à des températures relativement basses, plus pauvre que celle de la Terre. Ceci indique que la Lune a dû être soumise régulièrement à des températures plus élevées que la Terre, pendant toute une période de son histoire.

Supposons par exemple que la Lune ait commencé par être une planète, avec une orbite très excentrée, l'aphélie étant à peu près à la distance du Soleil correspondant à l'orbite de la Terre, et le périhélie à celle de Mercure. Quelques milliards d'années peuvent s'être passés ainsi, avant qu'une combinaison particulière des orbites de la Lune, de la Terre et peut-être de Vénus ait provoqué la capture de la Lune par la Terre. La Lune aurait ainsi cessé d'être une petite planète pour devenir un satellite, sa surface gardant la trace de ses voyages antérieurs dans les parages de Mercure.

Mais cette vitrification peut aussi bien résulter des échauffements locaux provoqués par le bombardement météorique qui a créé les cratères de la Lune. Ou dans l'hypothèse, hautement improbable, où la Lune aurait été arrachée à la Terre, la vitrification résulterait de la température élevée produite par un événement aussi violent.

En fait, toutes les suggestions relatives à l'origine de la Lune semblent également improbables, et on aurait entendu des astronomes grommeler que toute étude sérieuse des indices à ce sujet débouche forcément sur la conclusion que la Lune ne peut pas être là où elle est – ce qui veut dire qu'il faut chercher d'autres indices. Il y a forcément une réponse, et on finira par la trouver.

LA TERRE LIQUIDE

Puisque la Terre comporte deux parties bien distinctes, le manteau de silicates et le noyau de fer au nickel (avec une importance relative du même ordre que le blanc et le jaune d'un œuf), beaucoup de géologues pensent qu'elle a commencé par être liquide, et formée de deux liquides non miscibles. Le plus léger (le silicate) se serait rassemblé à la surface, et se serait refroidi en rayonnant dans l'espace, tandis que le fer central, isolé par cette couche externe du « froid » spatial, se serait refroidi beaucoup plus lentement, restant liquide jusqu'à notre époque.

Quant à savoir comment la Terre a pu s'échauffer au point de se liquéfier, à partir d'un groupe de planéticules froids, il y a au moins trois réponses possibles. D'abord, les chocs entre ces planéticules auraient transformé en

chaleur leur énergie de mouvement (*énergie cinétique*). Ensuite, à mesure qu'elle grossissait, la planète en formation aurait été comprimée par un champ de gravitation croissant, libérant encore de l'énergie sous forme de chaleur. Enfin, les corps radioactifs de la Terre (uranium, thorium et potassium) auraient libéré au fil des millénaires une énergie importante par leur fission, et comme ils étaient plus abondants au début, la radioactivité à elle seule aurait peut-être été capable de dégager assez de chaleur pour liquéfier la Terre.

Pourtant, il existe quelques scientifiques pour lesquels une étape liquide de la Terre n'apparaît pas comme une nécessité absolue. Par exemple, le chimiste américain Harold Clayton Urey soutient que la Terre a toujours été solide. Il affirme que, même dans une Terre « largement solide », un noyau de fer aurait pu tout de même se former lentement par lente séparation du fer ; et qu'actuellement, le fer pourrait bien être en train d'émigrer du manteau vers le noyau, à raison de 50 000 tonnes par seconde...

L'Océan

Chose unique parmi les planètes du système solaire, la température à la surface de la Terre permet à l'eau d'y exister sous ses trois états : glace, liquide, vapeur. La plupart des corps du système qui sont plus éloignés du Soleil que la Terre portent beaucoup de glace, Ganymède et Callisto par exemple. Quant à Europe, c'est un glacier à l'échelle d'une planète, avec peut-être de l'eau liquide en dessous. Mais tous ces mondes froids ne comportent que des quantités infimes de vapeur.

La Terre est le seul objet du système solaire – autant que nous sachions – à posséder des océans, c'est-à-dire des étendues d'eau liquide, en contact avec l'atmosphère qui les surmonte. Quand je dis « océans », d'ailleurs, je devrais dire *Océan*, parce que les océans Pacifique, Atlantique, Indien, Arctique et Antarctique forment une seule masse d'eau salée, dans laquelle l'Eurasie, l'Amérique et les blocs plus petits que sont l'Antarctique et l'Australie font figure d'îles.

Les caractéristiques numériques de cet Océan sont impressionnantes. Sa surface totale, 225 millions de kilomètres carrés, couvre 71 % de celle du globe. Son volume, en admettant une profondeur moyenne de 3 700 mètres, vaut 1,3 milliard de kilomètres cubes. Il contient 97,2 % de l'eau présente sur la Terre, et fournit aussi toute l'eau douce, puisque chaque année 320 000 kilomètres cubes de son eau s'évaporent, pour retomber sous forme de pluie ou de neige. En conséquence, il y a environ 800 000 kilomètres cubes d'eau douce sous la surface des continents, et 120 000 dans les rivières et les lacs.

Mais d'un autre point de vue, l'Océan est beaucoup moins impressionnant. Si grand soit-il, il ne forme que le quatre millième de la masse totale de la Terre. Si nous représentons celle-ci par une boule de billard, l'Océan ne serait qu'une couche imperceptible de buée. En descendant le plus loin possible au fond de l'Océan, on parcourrait moins des deux millièmes du rayon terrestre, le reste de la distance constitué d'abord de roches, puis de métal.

Mais pour nous, cette buée imperceptible fait la différence entre la vie et la mort. C'est là que les premières formes de vie se sont développées, et du simple point de vue de la masse vivante, les océans abritent encore la majorité de la vie sur notre planète. Sur la terre ferme, la vie ne s'élève qu'à un mètre ou deux de la surface – bien que les oiseaux et les avions fassent des sorties occasionnelles à de plus grandes hauteurs – tandis que dans l'Océan la vie occupe tout le volume, jusqu'à plus de dix kilomètres de profondeur par endroits.

Et pourtant, il y a quelques années, les hommes savaient aussi peu de choses sur les profondeurs de l'Océan, et en particulier sur son plancher, que si tout cela avait été situé sur la planète Vénus.

LES COURANTS

Le fondateur de l'océanographie moderne est un officier de marine américain nommé Matthew Fontaine Maury. A trente ans, un accident le rendit boiteux – accident pénible pour lui, mais qui devait se révéler bénéfique pour l'humanité. Nommé (par charité sans doute) responsable du dépôt des cartes et instruments, il s'attela à une tâche écrasante : relever les courants océaniques. En particulier, il étudia le Gulf Stream, sur lequel les premières recherches avaient été faites dès 1769 par l'érudit américain Benjamin Franklin. Maury en donna une description qui est devenue classique en océanographie : « Il y a une rivière dans l'Océan. » C'est une rivière beaucoup plus vaste que n'importe quelle rivière continentale. Son débit est mille fois supérieur à celui du Mississipi. Large de 80 kilomètres au départ, et profond de plus de 500 mètres, il se déplace à plus de 6 km/h. La tiédeur de ses eaux se fait sentir jusquau Spitzberg, aux confins de l'Arctique.

Maury fut aussi à l'origine de la coopération internationale en matière d'océanographie. Il fut le père de la conférence historique tenue à Bruxelles en 1853. En 1855, il publia le premier manuel d'océanographie, intitulé *Physical Geography of the Sea* (Géographie physique de la mer). L'Académie navale d'Annapolis a donné son nom au Maury Hall.

Depuis l'époque de Maury, on a relevé en détail le parcours des courants océaniques. Ils décrivent de grands cercles, dans le sens des aiguilles d'une montre dans l'hémisphère nord, et en sens inverse dans l'hémisphère sud, en raison de l'effet Coriolis. Le Gulf Stream n'est que la branche occidentale d'un courant circulaire (et direct) de l'Atlantique nord. Au sud de Terre-Neuve, ce courant vire plein est à travers l'Atlantique (*dérive nord-atlantique*). Les côtes européennes en dévient une partie vers la Norvège, en contournant les îles Britanniques. Le reste est dévié vers le sud, le long des côtes nord-ouest de l'Afrique. C'est le *courant des Canaries,* qui baigne l'archipel du même nom. La forme de la côte africaine s'ajoute à l'effet Coriolis pour renvoyer ce courant vers l'ouest, à travers l'Atlantique (*courant nord-équatorial*). Il arrive ainsi dans la mer des Caraïbes, et recommence le cercle.

Un tourbillon plus vaste, et dans le sens inverse des aiguilles d'une montre, suit les limites du Pacifique au sud de l'équateur. Le courant part de l'Antarctique, et remonte la côte américaine jusqu'au Pérou. Cette partie du parcours s'appelle *le courant froid du Pérou,* ou *courant de Humboldt* (du nom du naturaliste allemand Alexandre von Humboldt, qui l'a décrit le premier vers 1810).

Puis la forme de la côte péruvienne, s'ajoutant à l'effet Coriolis, dévie le courant vers l'ouest à travers le Pacifique, juste au sud de l'équateur (*courant sud-équatorial*). Une partie du courant se fraie un chemin entre les eaux de l'archipel indonésien et pénètre dans l'océan Indien. Le reste longe vers le sud la côte est de l'Australie, puis tourne encore vers l'est pour refermer sa boucle.

Ces mouvements d'eau contribuent à égaliser la température des océans, et indirectement celle des côtes, où il y a encore de grandes inégalités de température, mais beaucoup moins qu'il n'y en aurait en l'absence de courants océaniques.

La plupart des courants sont lents, plus lents encore que le Gulf Stream. Mais même à des vitesses aussi faibles, l'étendue considérable mise en jeu donne des valeurs énormes aux masses d'eau déplacées. Au large de New York, le Gulf Stream déplace vers le nord-est un débit de l'ordre de 45 millions de tonnes par seconde.

Il y a aussi des courants dans les régions polaires. Les grands courants circulaires (rétrogrades dans l'hémisphère nord et directs dans l'autre) ont tous pour effet, à leur limite polaire, de pousser l'eau d'ouest en est.

Au sud des continents américain, africain et australien, le courant tourne autour de l'Antarctique d'ouest en est suivant un cercle continu : c'est le seul endroit de la Terre où l'on peut décrire un parallèle tout entier sans rencontrer la côte. Ce courant antarctique est le plus important du monde en volume : chaque seconde, un méridien est traversé vers l'est par 100 millions de tonnes d'eau.

Dans la région arctique, le courant polaire est fractionné par les masses continentales en deux courants principaux : la *dérive nord-atlantique* et la *dérive nord-pacifique*. La première est déviée vers le sud par la côte du Groenland, et l'eau glaciale venue des pôles longe ainsi le Labrador et Terre-Neuve : c'est le *courant du Labrador*, qui rencontre le Gulf Stream au sud de Terre-Neuve, produisant dans cette région tempêtes et brouillards.

Le contraste est frappant entre les deux côtés de l'Atlantique. Le Labrador, exposé au courant qui porte son nom, est une désolation qui abrite à peine 25 000 personnes. De l'autre côté, exactement à la même latitude, les îles Britanniques en abritent 55 000 000, grâce au Gulf Stream.

Un courant qui se déplacerait en suivant exactement l'équateur ne serait pas dévié par effet Coriolis. Il en existe bien un, assez étroit, dans le Pacifique, qui suit l'équateur vers l'est sur plusieurs milliers de kilomètres. C'est le *courant de Cromwell*, du nom de celui qui l'a découvert, l'océanographe américain Townsend Cromwell. Un courant analogue, un peu moins important, existe aussi dans l'Atlantique, comme l'a montré en 1961 l'océanographe américain Arthur D. Voorhis.

D'autre part, cette circulation ne se limite pas à la surface. Différents résultats montrent clairement que les profondeurs de l'Océan ne peuvent pas être calmes. D'abord, la vie dans les couches supérieures consomme constamment leurs éléments nutritifs (nitrates et phosphates) et ces éléments descendent vers les profondeurs avec les cadavres de leurs consommateurs. S'il n'y avait aucune circulation capable de les faire remonter à la surface, celle-ci ne tarderait pas à en être dépourvue. Ensuite, l'oxygène absorbé par les océans, à partir de l'atmosphère, ne pourrait pas descendre assez profondément pour permettre la vie abyssale, s'il n'existait une circulation générale capable de l'entraîner. Or, sa concentration est

suffisante pour cela jusqu'au plancher le plus profond des fosses océaniques : la seule explication possible est qu'il y a des zones de l'Océan où l'eau chargée d'oxygène descend vers les profondeurs.

Le moteur de cette circulation verticale est la différence de température. Dans les régions polaires, l'eau de surface se refroidit et coule vers le fond. Ce courant continu s'étale en arrivant au fond, tout au long du plancher, de sorte que même sous les tropiques l'eau profonde est très froide (proche du point de congélation). Finalement, cette eau froide remonte vers la surface : elle n'a pas d'autre endroit où aller. En montant, elle se réchauffe et dérive vers les régions polaires, pour y couler à nouveau vers le fond. La circulation qui en résulte serait capable, si on ajoutait un nouveau produit en un point de l'océan Atlantique, de « mélanger » le tout en un millier d'années environ. Pour le Pacifique, plus grand, il faudrait peut-être deux mille ans.

L'Antarctique fournit beaucoup plus d'eau froide que l'Arctique : sa calotte glaciaire est dix fois plus étendue que toutes les glaces arctiques, y compris la calotte du Groenland. Tout autour de l'Antarctique, l'eau glacée (en contact avec la glace) se répand à la surface jusqu'à ce qu'elle rencontre l'eau tiède qui vient des tropiques. Plus dense que ces dernières, elle s'enfonce en dessous d'elles, sur la ligne de *convergence antarctique*, qui par endroits se rencontre jusqu'à des latitudes avoisinant 40 °S.

Cette eau froide de l'Antarctique s'étale sur le fond, emportant avec elle de l'oxygène (qui, comme tous les gaz, se dissout plus facilement dans l'eau froide que dans l'eau tiède) et des corps nutritifs. C'est ainsi que l'Antarctique (la « glacière du monde ») fertilise l'Océan et régule le climat de la planète.

Les barrières continentales compliquent ce modèle. Pour suivre la circulation générale, les océanographes suivent l'oxygène à la piste. Quand l'eau polaire, riche en oxygène, s'enfonce et s'étale, des êtres vivants se mettent à consommer cet oxygène, dont la concentration diminue donc à mesure que l'on s'éloigne de sa source. On peut ainsi repérer les directions des courants profonds.

Ces études montrent qu'un courant important passe de l'Arctique dans l'Atlantique, suivant le même parcours que le Gulf Stream, mais en profondeur et dans l'autre sens. Un autre passe de l'Antarctique à l'Atlantique sud. Le Pacifique, lui, ne reçoit rien de l'Arctique : le « goulot » constitué par le détroit de Behring est trop étroit et trop peu profond. Aussi le Pacifique nord est-il un cul-de-sac pour les eaux profondes, ce que révèle leur pauvreté en oxygène. Cela explique que des régions entières de ce vaste océan soient très peu peuplées, et jouent en mer le rôle des déserts continentaux. On constate la même chose dans les mers presque fermées, comme la Méditerranée, où l'oxygène et les éléments nutritifs circulent mal.

En 1957, une expédition océanographique anglo-américaine est parvenue à obtenir des résultats plus directs sur les courants profonds, grâce à un flotteur spécial, inventé par l'océanographe anglais John Crossley Swallow. Ce flotteur était conçu pour se maintenir à une profondeur constante, pouvant aller jusqu'à deux kilomètres, et muni d'un émetteur d'ultrasons. Le signal émis permettait de suivre le flotteur tandis qu'il était entraîné par les courants profonds. L'expédition a pu de cette manière dresser la carte du courant qui descend le long de la côte ouest de l'Atlantique.

LES RESSOURCES DE L'OCÉAN

Toutes ces informations prendront une grande importance pratique le jour où les hommes, de plus en plus nombreux, devront chercher dans l'Océan une partie de leur subsistance. Pour exploiter scientifiquement la mer, il faudra connaître ces courants fertilisateurs, comme l'agriculteur connaît les rivières et le régime des pluies. La production actuelle des pêcheries – quelque quatre-vingts millions de tonnes en 1980 – pourrait être portée, grâce à une organisation soigneuse, jusqu'à plus de deux cents millions de tonnes par an, tout en permettant à la vie marine de maintenir ses effectifs. Ceci suppose, bien entendu, que nous cessions d'abîmer et de polluer l'Océan, surtout les régions – près des côtes continentales – qui contiennent la plus grande partie de la faune. Pour l'instant, non seulement nous n'arrivons pas à rationaliser l'exploitation de la mer pour en tirer davantage parti, mais nous diminuons sa capacité à nous fournir ce que nous en tirons actuellement.

La nourriture n'est pas la seule ressource importante de l'Océan. L'eau de mer contient en solution de grandes quantités de la plupart des éléments. On estime à quatre milliards de tonnes la masse d'uranium qu'elle contient, trois cents millions de tonnes celle de l'argent et à quatre millions de tonnes celle de l'or – tout cela à des dilutions trop faibles pour permettre pratiquement l'extraction de ces métaux. Mais la production de brome et de magnésium à partir de l'eau de mer, elle, est déjà industrielle. De plus, les algues séchées fournissent une source importante d'iode, ces plantes étant plus efficaces que les procédés dont nous disposons pour augmenter la concentration de cet élément à partir de ce que contient l'eau de mer.

On tire aussi de la mer des matériaux plus prosaïques. Des eaux plus profondes qui bordent les États-Unis, on tire chaque année 20 millions de tonnes de coquilles d'huîtres, qui fournissent une quantité appréciable de chaux. Dans le même temps, la récolte de sable et de gravier atteint 50 millions de mètres cubes.

Les régions les plus profondes du plancher océanique sont parsemées de nodules métalliques, qui se sont formés par précipitation autour d'un galet ou d'une dent de requin (comme une perle se forme autour d'un grain de sable à l'intérieur de l'huître). On les appelle habituellement « nodules de manganèse », parce que c'est le métal qui forme la plus grande partie de leur masse. On estime qu'un kilomètre carré au fond du Pacifique porte plus de dix mille tonnes de ces nodules. Ils seraient difficiles à récolter, bien sûr, et leur contenu en manganèse ne suffirait pas, actuellement, à rendre l'opération rentable. Mais les nodules contiennent aussi 1 % de nickel, 0,5 % de cuivre et 0,5 % de cobalt, ce qui augmente considérablement leur intérêt.

Enfin, l'Océan contient autre chose que des matériaux dissous : de l'eau (97 % de sa masse).

Les Américains consomment 3 000 mètres cubes d'eau par an et par personne pour boire, pour se laver, pour les autres usages domestiques, agricoles et industriels. La plupart des autres pays montrent davantage de modération, mais la moyenne mondiale s'élève quand même à 1 500 mètres cubes par personne et par an. De plus, il s'agit d'*eau douce.* L'eau de mer, telle qu'elle est, est impropre à tous ces usages.

La Terre porte une grande quantité d'eau douce. Bien que 3 % seulement de toute l'eau présente sur Terre soit de l'eau douce, cela fait tout de même

dix millions de mètres cubes par personne. Toutefois, les trois quarts de cette eau sont inutilisables, étant stockés dans les calottes glaciaires permanentes qui couvrent 10 % de la terre ferme.

Finalement, l'eau douce liquide disponible sur Terre représente trois millions de mètres cubes par personne, et se renouvelle constamment grâce à la pluie, à la cadence de plus de cent mille mètres cubes par personne et par an. Comme c'est à peu près soixante-quinze fois la consommation moyenne, on pourrait croire qu'il n'y a pas de problème.

Seulement, la plus grande partie de la pluie tombe sur l'Océan, ou – sous forme de neige – sur les calottes glaciaires. De celle qui tombe sur la terre, et qui reste liquide – ou le devient en se réchauffant – une bonne partie retourne à la mer sans être utilisée. Une autre partie est pratiquement inutilisable dans le bassin de l'Amazone. La population humaine s'accroît constamment, et de plus pollue sans relâche les réserves d'eau douce existantes.

Aussi, d'ici peu, on manquera d'eau douce, et on commence déjà à se tourner ver l'Océan. On peut distiller l'eau de mer, en recueillant par condensation l'eau évaporée, les matériaux dissous restant sous forme solide. Idéalement, on peut utiliser à cet effet l'énergie du Soleil. Les méthodes de *désalinisation* peuvent fournir de l'eau douce, là où le soleil est généreux, le fuel bon marché, ou les besoins impérieux. Par exemple, les grands paquebots utilisent une partie de leur carburant pour dessaler l'eau dont ils ont besoin.

Enfin, on commence à étudier la possibilité de remorquer des icebergs depuis des régions polaires jusqu'à des ports chauds et secs, où la glace qui aurait survécu au voyage pourrait fournir de l'eau douce.

Mais sans aucun doute, le meilleur moyen d'utiliser nos ressources en eau douce (ou d'ailleurs n'importe quelles ressources) est de les ménager soigneusement, de réduire au maximum le gaspillage et la pollution, et de limiter prudemment la population humaine.

PROFONDEURS OCÉANIQUES ET ÉVOLUTION DES CONTINENTS

Qu'en est-il de l'observation directe des profondeurs de l'Océan ? L'Antiquité nous a laissé sur ce point un seul résultat (auquel il ne faut peut-être pas accorder une confiance totale) : le philosophe grec Posidonius, environ en l'an 100 av. J.-C., aurait mesuré la profondeur de la Méditerranée près des côtes de Sardaigne, et aurait relevé un fond de 1 900 mètres environ.

Mais l'étude systématique des fonds, dans un but zoologique, n'a vraiment commencé qu'au XVIIIe siècle. Vers 1770, le biologiste danois Otto Frederik Muller a mis au point une drague permettant de rapporter des profondeurs des échantillons de la vie marine.

Une drague de ce type allait permettre à un biologiste anglais, Edward Forbes Jr., d'obtenir des résultats importants. Vers 1830, il recueillit ainsi des échantillons dans la mer du Nord et dans toutes les eaux qui baignent les îles Britanniques. Puis, en 1841, il embarqua sur un navire de guerre à destination de la Méditerranée orientale, d'où il rapporta une étoile de mer draguée à 450 mètres de profondeur.

Or, la lumière du Soleil ne pénètre que jusqu'à 80 mètres, et donc, plus bas, il n'y a pas de plantes. Les animaux se nourrissant de plantes, Forbes

pensa qu'ils ne pouvaient pas rester longtemps à un niveau où les plantes ne poussaient plus. Il évalua à 450 mètres de profondeur la limite de la vie, les couches plus profondes étant désertes et stériles.

Pourtant, juste au moment où Forbes publiait sa théorie, l'explorateur anglais James Clark Ross, dans les parages de l'Antarctique, draguait à 800 mètres de fond des échantillons vivants – bien au-dessous de la limite assignée par Forbes. Mais l'Antarctique est loin, et la plupart des biologistes firent confiance à Forbes.

Au milieu du XIXe siècle, le fond des mers allait acquérir un intérêt pratique pour tout le monde (il n'avait jusque-là qu'un intérêt scientifique pour quelques spécialistes), avec les projets de câbles transatlantiques. A cet effet, Maury établit en 1850 une carte du fond de l'Atlantique. Il fallut quinze ans, marqués par de nombreux échecs, pour que le câble finisse par être posé – sous l'impulsion inlassable du financier américain Cyrus West Field, qui y consacra une fortune. Maintenant, plus de vingt câbles traversent l'Atlantique.

A cette occasion, et grâce à Maury, l'exploration systématique du fond océanique a commencé. Les sondages de Maury ont révélé que l'Atlantique était moins profond au centre que des deux côtés. Maury a donné à ce bas-fond central le nom de plateau du Télégraphe, en l'honneur du câble.

Le navire anglais *Bulldog* allait poursuivre et étendre les explorations des fonds marins entreprises par Maury. Il partit en 1860, emmenant à son bord un médecin anglais, George C. Wallich. Celui-ci parvint, avec une drague, à remonter treize étoiles de mer d'une profondeur de 2 500 mètres. Et ce n'étaient pas des cadavres qui avaient coulé à pic : elles étaient bien vivantes. Aussitôt, Wallich annonça sa découverte, qui montra que la vie animale pouvait exister dans l'obscurité glaciale des grandes profondeurs, même sans aucune végétation.

Pourtant, les biologistes hésitaient encore à admettre cette possibilité. Mais en 1868 un biologiste écossais, Charles W. Thomson, draguant le fond à partir d'un navire appelé *Lightning* (l'éclair), recueillit de multiples échantillons animaux, mettant fin à la discussion, et donnant le coup de grâce à l'idée de Forbes suivant laquelle la vie était limitée aux couches peu profondes de l'Océan.

Voulant déterminer la profondeur de l'Océan, Thomson s'embarqua le 7 décembre 1872 à bord du *Challenger*. Il allait passer trois ans et demi en mer, et parcourir plus de 120 000 kilomètres. Pour mesurer la profondeur de l'Océan, le *Challenger* ne disposait que de la méthode séculaire du sondage : laisser filer un câble de six kilomètres, lesté par un poids, jusqu'à ce que celui-ci touche le fond. Thomson réalisa 370 sondages par cette méthode. Malheureusement celle-ci n'est pas seulement très laborieuse (dès que la profondeur devient importante), elle est aussi très peu précise. C'est seulement en 1922 que tout changea, et que le *sondage par écho* au moyen d'ondes sonores révolutionna l'exploration des fonds marins. Pour expliquer le principe de cette méthode, une digression sur l'acoustique s'impose.

Des vibrations mécaniques peuvent établir dans un milieu matériel (l'air par exemple) des ondes longitudinales, dont certaines sont capables d'impressionner notre oreille, leurs différentes longeurs d'onde correspondant à des sons de hauteurs différentes. Le son le plus grave que nous soyons capables d'entendre a une longueur d'onde de 22 mètres, et une fréquence de 15 cycles par seconde. Le son le plus aigu perceptible pour

un adulte normal a une longueur d'onde de 2,2 centimètres, et une fréquence de 15 000 cycles par seconde (les enfants perçoivent des sons un peu plus aigus).

L'absorption du son par l'atmosphère dépend de la longueur d'onde. Plus celle-ci est grande, moins le son est absorbé par une épaisseur donnée d'air. C'est pourquoi les sirènes de brume ont un son très grave, de manière à être entendues le plus loin possible. Celles d'un grand paquebot comme le vieux *Queen Mary* ont une fréquence de 27 vibrations par seconde, à peu près celle de la note la plus grave d'un piano. On l'entend à 16 kilomètres, et des instruments convenables peuvent la détecter à 200 kilomètres.

Il existe aussi des « sons » plus graves que le plus grave que nous soyons capables d'entendre. Par exemple, les tremblements de terre et les volcans produisent – entre autres sons – des vibrations dans cette gamme *infrasonique.* Certaines de ces ondes peuvent faire le tour de la Terre, parfois à plusieurs reprises, avant d'être complètement absorbées.

Quant à la faculté du son de se réfléchir sur les obstacles, elle dépend aussi de la longueur d'onde, mais en sens inverse : plus la longueur d'onde est petite, mieux le son se réfléchit. L'efficacité est encore plus grande avec les *ultrasons* trop aigus pour que nous puissions les percevoir. Certains animaux pourtant les perçoivent, et se servent de cette capacité. Les chauves-souris émettent des « cris » ultrasoniques, dont la fréquence va jusqu'à 130 000 cycles par seconde, et elles écoutent les échos qui leur reviennent. A partir de la direction de ces échos, et du temps au bout duquel ils arrivent, l'animal estime la position de l'insecte à capturer ou du rameau à éviter. Une chauve-souris vole parfaitement si on l'aveugle, pas si on la rend sourde (l'Italien Lazzaro Spallanzani, qui fit le premier cette observation en 1793, se demandait si les chauves-souris « voyaient avec leurs oreilles », et en un sens il n'avait pas tort).

La même méthode de *localisation par écho* est utilisée par les guacharos (oiseaux des cavernes du Venezuela) et surtout par les dauphins. Ceux-ci, devant localiser des objets plus gros que les chauves-souris, peuvent se permettre d'utiliser des sons moins efficaces, à la limite de la gamme audible. On se demande même si les sons complexes émis par les dauphins ne leur servent pas aussi à communiquer entre eux – à parler, en somme... Les recherches exhaustives menées sur ce point par le biologiste américain John C. Lilly n'ont pas donné de résultats concluants.

Pour utiliser les ultrasons, les hommes doivent commencer par les produire. A petite échelle, cette production et cette utilisation trouvent un exemple dans le *sifflet à chien* (inventé en 1883). Cet instrument émet des sons dans la gamme ultrasonique proche, inaudibles pour l'homme, mais parfaitement audibles pour le chien.

Un chemin beaucoup plus fructueux allait s'ouvrir en 1880. Cette année-là, les frères Curie (Pierre et Jacques) découvrirent la *piézo-électricité :* une pression exercée sur certains cristaux produit un potentiel électrique, et réciproquement. En appliquant une différence de potentiel électrique à un cristal de ce type, il se contracte légèrement comme s'il était soumis à une pression mécanique (*électrostriction*). Il suffit de produire un potentiel alternatif de fréquence suffisamment élevée pour donner au cristal des vibrations de même fréquence, et émettre ainsi des ultrasons. C'est ce qu'a fait, dès 1917, le physicien français Paul Langevin, qui a appliqué immédiatement à la détection des sous-marins la réflexion excellente

obtenue avec ces ondes. La Première Guerre mondiale s'est terminée, mais lors de la Deuxième, le procédé s'est perfectionné, et il est devenu le *sonar* (*so*und *n*avigation *a*nd *r*anging, c'est-à-dire « navigation et télémétrie à partir du son »).

Pour déterminer la profondeur de la mer, la réflexion des ultrasons remplace la ligne de sonde. L'intervalle de temps entre l'émission du signal (une impulsion très brève) et le retour de son écho permet de calculer facilement la distance du fond. L'opérateur doit seulement se méfier des « faux échos » produits par des obstacles sur le chemin, des bancs de poissons par exemple (ce qui indique par ailleurs l'utilité de l'instrument pour les pêcheurs).

Non seulement la méthode est rapide et commode mais elle permet de travailler d'une façon continue, et donne aux océanographes le tracé du fond au-dessus duquel leur navire se déplace. On obtient ainsi en cinq minutes plus de détails que le *Challenger* n'aurait pu en accumuler pendant tout son voyage.

Le premier navire à utiliser le sonar de cette manière a été le *Meteor*, navire océanographique allemand qui a étudié l'océan Atlantique en 1922. Dès 1925, il est devenu évident que le fond de l'Océan n'était pas du tout une plaine uniforme, et que le plateau du Télégraphe n'était pas une douce ondulation : c'est une chaîne de montagnes, plus longue et plus chaotique que n'importe quelle chaîne terrestre. Elle court tout au long de l'Atlantique, et ses pics les plus élevés émergent, sous la forme des Açores, des îles de l'Ascension et de Tristan da Cunha. On lui donne le nom de chaîne centrale de l'Atlantique.

Au fil des années, les découvertes spectaculaires se sont multipliées. L'île de Hawaï est le sommet d'une montagne sous-marine de plus de 10 000 mètres (à partir du fond) – plus que n'importe quel pic de l'Himalaya : on peut dire que Hawaï est la plus haute montagne du monde. Les fonds marins portent aussi des montagnes curieuses, en forme de troncs de cônes, appelées *guyots* en l'honneur du géographe américano-suisse Arnold Henry Guyot, qui a introduit la géographie scientifique aux État-Unis en venant s'installer dans ce pays en 1848. Les premiers guyots ont été découverts pendant la Seconde Guerre mondiale par le géologue américain Harry Hammond Hess, qui en a trouvé dix-neuf coup sur coup. Il y en a au moins 10 000 , surtout dans le Pacifique. L'un d'eux, découvert en 1964 au sud de l'île de Wake, a plus de 4 000 mètres de haut.

Autre relief spectaculaire du fond des mers : les fosses abyssales, de plus de 6 000 mètres de profondeur, dans lesquelles le Grand Canyon passerait inaperçu. Toutes situées auprès d'un archipel, elles ont une surface totale avoisinant 1 % de la surface totale des océans. Cela paraît peu, mais c'est tout de même la moitié de la surface des États-Unis, et les fosses contiennent plus d'eau que toutes les rivières et les lacs de la Terre. Les plus profondes sont dans le Pacifique, près des archipels des Philippines, des Mariannes, des Kouriles, des îles Salomon et des Aléoutiennes (figure 4.5). Il y a aussi des fosses dans l'Atlantique, près des Antilles, et une dans l'océan Indien.

On trouve également au fond de l'Océan des gorges, parfois longues de milliers de kilomètres, et qui ressemblent aux vallées des rivières terrestres. Certaines d'entre elles semblent d'ailleurs prolonger des rivières continentales, en particulier la gorge qui s'étend à travers l'Atlantique en prolongeant l'Hudson. Dans la seule baie du Bengale, on a relevé vingt de ces gorges, lors des études océanographiques de l'océan Indien dans

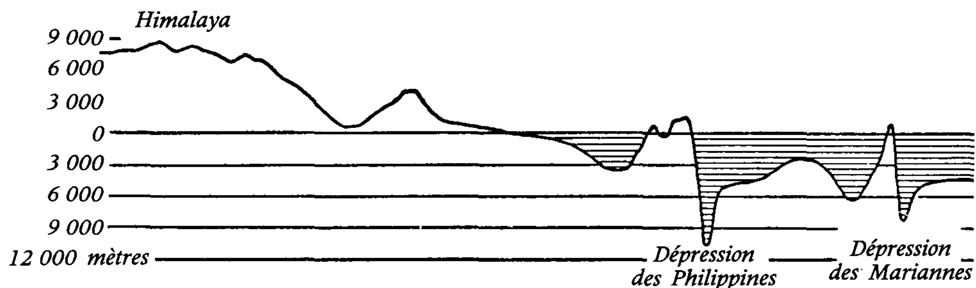
les années soixante. Il est tentant de voir là les lits de rivières situées autrefois sur la terre ferme, quand l'Océan était moins profond. Mais certaines de ces gorges sont tellement au-dessous du niveau actuel de la mer qu'il semble tout à fait invraisemblable qu'elles aient pu un jour être sur la terre ferme. Depuis quelques années, plusieurs océanographes – notamment William Maurice Ewing et Bruce Charles Heezen – ont avancé une autre théorie : les gorges auraient été creusées par des courants turbulents d'eau limoneuse dévalant les contreforts des continents à des vitesses pouvant atteindre 100 km/h. Un de ces courants turbulents – qui a attiré l'attention du monde scientifique sur le phénomène – a été mis en branle en 1929 par un tremblement de terre près de Terre-Neuve. Il a coupé l'un après l'autre plusieurs câbles sous-marins, et a causé des dégâts considérables.

La chaîne centrale atlantique tenait d'autres surprises en réserve. Des sondages ultérieurs ont montré qu'elle ne se limitait pas à l'Atlantique. A son extrémité méridionale, elle contourne l'Afrique, et traverse l'océan Indien jusqu'à l'Arabie. Au milieu de l'océan Indien, elle bifurque et continue au sud de l'Australie et de la Nouvelle-Zélande, pour se prolonger ensuite vers le nord à travers tout le Pacifique. Plus que d'une chaîne centrale atlantique, il s'agit d'une chaîne centrale océanique. Et cette chaîne diffère fondamentalement des chaînes de montagnes terrestres : celles-ci sont formées par les plissements de couches sédimentaires, tandis que la grande chaîne océanique est faite du basalte qui monte des profondeurs chaudes de la croûte terrestre.

Après la Deuxième Guerre mondiale, Ewing et Heezen ont exploré avec une ardeur nouvelle les détails du plancher océanique. En 1953 ils ont découvert, à leur grande surprise, qu'une faille profonde partageait en deux, sur toute sa longueur, la chaîne centrale. On a constaté ensuite que cette faille était présente sur toutes les parties de la chaîne, dans tous les océans. A certains endroits elle s'approche de la Terre : elle suit la mer Rouge entre l'Afrique et l'Arabie, et longe la côte californienne.

On a d'abord cru qu'il s'agissait d'une faille continue, une fêlure de 70 000 kilomètres dans l'écorce terrestre. Mais en y regardant de plus près,

Figure 4.5. Profil du fond de l'océan Pacifique. Les grandes fosses ont une profondeur supérieure à la hauteur de l'Himalaya, et le pic de Hawaï s'élève au-dessus du fond plus que les plus hautes montagnes au-dessus du niveau de la mer.



on s'est aperçu qu'elle était constituée de petites sections rectilignes décalées les unes par rapport aux autres, comme si des tremblements de terre les avaient cisaillées. Et de fait, c'est tout au long de cette faille que les tremblements de terre et les volcans se multiplient.

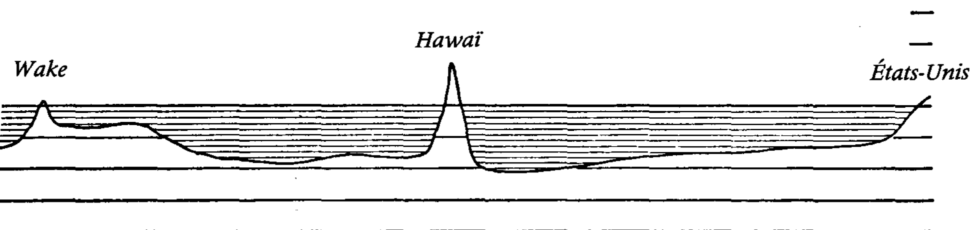
C'est un point faible, à travers lequel la roche en fusion (le *magma*) suinte lentement, s'empilant à mesure qu'il se refroidit pour former la chaîne, et continuant à s'étendre des deux côtés. Cet étalement peut atteindre une vitesse de seize centimètres par an – ce qui couvrirait le Pacifique tout entier en cent millions d'années. D'ailleurs, les sédiments que l'on peut tirer du plancher océanique sont rarement plus vieux que cela, ce qui serait étrange sur une planète quarante-cinq fois plus ancienne, mais s'explique fort bien par l'idée d'étalement du plancher.

Il apparaissait donc que la croûte terrestre était divisée en grandes plaques, séparées les unes des autres par la grande faille et ses diverses branches. On leur a donné le nom de *plaques tectoniques* (d'un mot grec évoquant l'assemblage de pièces pour former un tout bien ajusté). On appelle *tectonique des plaques* l'étude de l'évolution de la croûte terrestre fondée sur l'existence et les mouvements de ces plaques.

Il y en a six grandes et un certain nombre de petites, et on a remarqué que les tremblements de terre se produisaient en général à leurs limites. Les limites de la plaque du Pacifique (qui couvre presque tout l'océan Pacifique) comprennent ainsi les zones sismiques d'Indonésie, du Japon, de l'Alaska, de la Californie, etc. En Méditerranée, la frontière entre la plaque eurasienne et la plaque africaine ne le cède qu'à la précédente en matière d'activité sismique.

D'autre part, les *failles* détectées dans la croûte terrestre, profondes fractures dont un côté glisse de temps en temps le long de l'autre, provoquant un tremblement de terre, étaient aussi le long des limites des plaques. La plus célèbre de ces failles, celle de San Andreas, qui court tout au long de la Californie de San Francisco à Los Angeles, fait partie de la limite entre la plaque américaine et la plaque pacifique.

Qu'en est-il de la dérive des continents, suivant la théorie de Wegener ? Si on considère une plaque, les objets qu'elle porte ne peuvent pas changer de position. Ils sont fixés en place par la rigidité du basalte, chère aux



adversaires de Wegener. De plus, les plaques voisines s'emboîtent si étroitement qu'on voit mal comment elles pourraient bouger.

C'est une autre constatation qui allait fournir la solution de ce problème. Sur les limites des plaques, non seulement les tremblements de terre sont fréquents, mais les volcans aussi. Les bords du Pacifique, tout le long des limites de la plaque du même nom, portent tant de volcans – actifs ou non – qu'on appelle l'ensemble le *cercle de feu*.

Se pourrait-il donc que le magma s'élève des couches chaudes très profondes à travers les fissures qui séparent les plaques, fissures qui représentent les seuls points faibles de la croûte terrestre ? Par exemple, le magma pourrait sortir lentement à travers la faille centrale atlantique, et se solidifier au contact de l'eau pour former, de part et d'autre, la chaîne centrale.

Mieux : en se solidifiant, le magma peut écarter les plaques. S'il en est ainsi, il serait responsable de la séparation de l'Afrique et de l'Amérique du Sud, de l'Europe et de l'Amérique du Nord, bref de la rupture de la Pangée et de la formation de l'Atlantique, puis de son élargissement progressif. De la même façon, l'Europe et l'Afrique s'écarteraient aussi, la Méditerranée s'élargissant entre les deux. Cette explication, appelée *élargissement du plancher océanique*, est due à H.H. Hess et Robert S. Dietz, qui l'ont proposée en 1960. Les continents ne dérivent pas, comme le pensait Wegener : ils sont fixés à des plaques qui sont lentement écartées les unes des autres.

Peut-on trouver des preuves de cet élargissement du plancher océanique ? A partir de 1963, on a mesuré le magnétisme d'échantillons rocheux recueillis de part et d'autre de la faille centrale de l'Atlantique. Ce magnétisme varie d'une façon symétrique des deux côtés, ce qui permet d'affirmer que les roches sont très jeunes près de la faille, et de plus en plus vieilles à mesure qu'on s'en éloigne, d'un côté et de l'autre.

On a pu ainsi estimer que l'Atlantique s'élargissait actuellement de 2,5 centimètres par an, et donc déterminer grossièrement l'époque à laquelle il a commencé à se former. Sur ce point et sur bien d'autres la tectonique des plaques, en vingt ans, a complètement révolutionné la géologie.

Evidemment, si deux plaques se séparent, chacune doit, de l'autre côté, se trouver comprimée contre sa voisine : l'ensemble est bien ajusté. Quand deux plaques se rencontrent ainsi, à des vitesses de l'ordre de quelques centimètres par an, la croûte se déforme et se plisse, à la fois au-dessus et au-dessous, formant des montagnes et leurs « racines ». C'est ainsi que l'Himalaya a sans doute été formé par l'écrasement continu de la plaque indienne contre celle qui porte le reste de l'Asie.

Quand le rapprochement est trop rapide pour permettre cette déformation continue de la croûte, l'une des plaques peut plonger sous l'autre, formant une tranchée profonde, une ligne d'îles, et une région à haute activité volcanique. C'est ce qui se passe par exemple à la limite occidentale du Pacifique, célèbre pour ses fosses abyssales, ses chapelets d'îles et ses volcans.

Aussi bien que là où elles se rapprochent, on peut trouver des preuves du mouvement des plaques là où elles s'écartent, sous l'effet de l'élargissement du plancher. La grande faille traverse l'Islande occidentale, qui s'élargit (très lentement). La mer Rouge est de formation très récente. Elle est apparue à la séparation de l'Afrique et de l'Arabie, et ses côtes

opposées « s'emboîtent » parfaitement l'une dans l'autre. Le processus continue, ce qui fait en somme de la mer Rouge un nouvel océan en formation. Or, on peut y trouver des preuves directes de la montée du magma : au fond de cette mer, on a trouvé, à partir de 1965, des zones où la température atteint 56 °C, et où la concentration saline est plus de cinq fois supérieure à la normale.

Vraisemblablement, une évolution très longue, très lente, voit le magma surgir, écarter les plaques, ce qui les rapproche en d'autres endroits, où la croûte plonge profondément et redevient du magma. Les continents se soudent, forment un seul bloc continental, qui se fracture, sans doute à de nombreuses reprises. Des montagnes se forment, s'usent et disparaissent, des fosses océaniques se creusent et se comblent, des volcans naissent et s'éteignent. La Terre n'est pas seulement le siège de la vie au sens biologique du terme ; géologiquement aussi, elle est vivante.

Les géologues sont même maintenant capables de suivre le cours de la dernière rupture de la Pangée, d'une façon encore assez grossière. La première fracture suit une ligne est-ouest. La moitié nord de la Pangée (ce qui forme maintenant l'Asie, l'Europe et l'Amérique du Nord) est souvent appelée Laurasie, parce que la partie la plus ancienne, géologiquement, des roches superficielles d'Amérique du Nord, forme les hauteurs qui dominent la rive nord du Saint-Laurent.

La moitié sud de la Pangée comprend tout ce qui forme actuellement l'Amérique du Sud, l'Afrique, l'Inde, l'Australie et l'Antarctique. On l'appelle Gondwana (nom proposé vers 1890 par le géologue autrichien Edward Suess, à partir de celui d'une province indienne, et sur la base d'une théorie de l'évolution géologique qui semblait raisonnable à l'époque, mais s'est, depuis, révélée fausse).

Il y a environ 200 millions d'années, l'Eurasie et l'Amérique du Nord se séparaient et dérivaient – pardon, étaient poussées – vers le nord, jusqu'à embrasser entre elles les régions arctiques. Cinquante millions d'années plus tard, l'Afrique et l'Amérique du Sud se séparaient et commençaient à s'éloigner l'une de l'autre. La seconde a fini par se rapprocher de l'Amérique du Nord, les deux continents se rejoignant suivant l'étroite bande de l'Amérique centrale.

Il y a environ 110 millions d'années, la partie orientale du Gondwana se fragmentait en plusieurs morceaux : Madagascar, l'Inde, l'Australie et l'Antarctique. Madagascar ne s'est pas beaucoup éloignée de l'Afrique, mais l'Inde a voyagé plus qu'aucun autre morceau de la Pangée la plus récente : elle a parcouru près de 9 000 kilomètres vers le nord, jusqu'à rencontrer l'Eurasie et y faire surgir l'Himalaya, le Pamir et le plateau tibétain – la région haute la plus jeune, la plus grande et la plus impressionnante de la Terre.

L'Antarctique et l'Australie se sont peut-être séparées seulement il y a quarante millions d'années, la première partant vers le sud, et le destin glacial qu'elle devait y connaître, la seconde vers le nord – voyage qu'elle poursuit encore aujourd'hui.

LA VIE DANS LES PROFONDEURS

Après la Seconde Guerre mondiale, l'exploration des profondeurs océaniques a continué. Depuis peu on a mis au point un « micro » sous-marin, l'*hydrophone,* qui révèle les bruits émis par les créatures

marines, cliquetis, grognements, claquements, gémissements, qui font dans les profondeurs un tapage aussi infernal que celui que nous connaissons sur la terre ferme.

Un nouveau *Challenger* a exploré en 1951 la fosse des Mariannes, dans l'ouest du Pacifique, et constaté que c'était elle, et non pas la fosse des Philippines, qui détenait le record de profondeur dans la croûte terrestre. La partie la plus profonde, baptisée *gouffre du Challenger,* a plus de onze kilomètres de profondeur. L'Everest pourrait y tenir à l'aise, en laissant deux kilomètres d'eau au-dessus de son sommet. Or, du fond de cet abîme, le *Challenger* a rapporté des bactéries qui ressemblaient beaucoup à celles de la surface – mais elles ont besoin pour vivre d'une pression au moins égale à mille atmosphères !

En effet, les habitants de ces fosses sont tellement adaptés aux pressions énormes qui y règnent qu'ils sont incapables d'en sortir : ils sont « prisonniers sur une île ». C'est le résultat d'une évolution séparée. Pourtant, sous bien des aspects, ils sont si proches d'autres êtres vivants que leur adaptation à l'abîme ne doit pas être très ancienne. On peut se représenter des groupes d'animaux marins poussés à des profondeurs croissantes par la compétition avec d'autres, exactement comme d'autres groupes ont été poussés de plus en plus haut le long des contreforts océaniques, jusqu'à devoir se hisser sur la terre ferme. Les premiers ont dû s'adapter à des pressions plus fortes, les seconds à l'absence d'eau : ceci était sans doute plus difficile que cela, et nous ne devrions pas être étonnés de trouver de la vie dans les abysses.

Evidemment, la vie n'est pas aussi riche dans les profondeurs que plus près de la surface. La masse de matière vivante à 7 000 mètres de profondeur est seulement le dixième de ce qu'elle est à 3 000 mètres. De plus, aux grandes profondeurs, il y a peu de carnivores, car ils n'auraient pas assez de proies à leur disposition. On y trouve des charognards, qui mangent tous les débris organiques existants. Cette colonisation des profondeurs est très récente : aucune des espèces n'a plus de 200 millions d'années, et la plupart datent de moins de 50 millions d'années. C'est seulement au début du règne des dinosaures qu'a commencé le peuplement des profondeurs, jusque-là désertes.

Néanmoins, cette émigration vers le fond a permis à certaines espèces de survivre tandis que leurs « parents » des zones superficielles s'éteignaient, comme on a pu le constater de façon spectaculaire à la fin des années trente. Le 25 décembre 1938, au large de l'Afrique du Sud, un chalutier a ramené un poisson bizarre, long environ de 1,50 mètre. Sa principale bizarrerie résidait dans ses nageoires, qui n'étaient pas fixées directement sur son corps, mais sur des « membres » distincts. Par chance, un zoologiste sud-africain, J.L.B. Smith, l'a examiné : ce fut pour lui le plus beau des cadeaux de Noël. Il s'agissait en effet d'un spécimen vivant d'un animal que l'on croyait disparu depuis 70 millions d'années, un *cœlacanthe,* poisson primitif éteint – en principe – avant même l'apogée des dinosaures.

La Seconde Guerre mondiale retarda la chasse au cœlacanthe, mais en 1952 on en a pêché un autre, d'espèce différente, près de Madagascar. Depuis, on en a trouvé un certain nombre. Adapté à des profondeurs importantes, le cœlacanthe meurt très vite si on le remonte à la surface.

Pour les évolutionnistes, l'étude du cœlacanthe est fascinante : c'est à partir de lui que se sont développés les amphibiens. En d'autres termes,

les cœlacanthes actuels descendent directement de nos ancêtres des bas-fonds...

A la fin des années soixante-dix, on a fait une découverte encore plus étonnante. Il y a dans l'Océan des *points chauds* où le magma brûlant du manteau se rapproche d'une façon inhabituelle de la limite supérieure de la croûte, et chauffe l'eau qui la surmonte.

L'exploration détaillée de ces points chauds a commencé en 1977, avec des sous-marins capables de descendre à de grandes profondeurs. On a étudié ainsi des points chauds près des Galapagos et à l'entrée du golfe de Californie. Là, on a trouvé des *cheminées* crachant des bouffées chaudes de boues fumantes, qui enrichissent en minéraux l'eau avoisinante.

Ces minéraux sont riches en soufre, et autour de ces points chauds on rencontre diverses espèces de bactéries, qui tirent leur énergie de réactions chimiques utilisant du soufre et de la chaleur, au lieu de la tirer du rayonnement solaire. De petits animaux mangent ces bactéries, et des animaux plus gros mangent les petits.

Voilà donc une chaîne d'êtres vivants tout entière, absolument indépendante de la vie végétale qui se développe dans les couches supérieures de la mer. Même s'il n'existait pas de rayonnement solaire, cette chaîne pourrait subsister, à condition que chaleur et éléments minéraux continuent à sortir des profondeurs terrestres, c'est-à-dire au voisinage des points chauds.

On a pu recueillir dans ces régions des coquillages, des crabes et diverses sortes de vers, dont certains de belle taille. Tous ces animaux vivent gaillardement dans une eau qui empoisonnerait n'importe quelle espèce qui ne serait pas spécialement adaptée aux particularités chimiques locales.

LA PLONGÉE PROFONDE

Ce dernier exemple montre l'intérêt d'envoyer des équipes humaines observer sur place les profondeurs marines. Bien sûr, c'est un milieu qui ne nous convient pas beaucoup. Dès l'Antiquité, les plongeurs sont parvenus, avec de l'entraînement, à descendre à une vingtaine de mètres et à rester sous l'eau jusqu'à deux minutes. Sans appareils, on ne peut guère améliorer ces performances.

Dans les années trente, les lunettes de plongée, les palmes et les *schnorkels* (tuyaux courts dont une extrémité est dans la bouche et l'autre dépasse à la surface de l'eau) augmentent l'efficacité des plongeurs, et le temps qu'ils peuvent passer sous l'eau.

En 1943, un officier de marine français, Jacques-Yves Cousteau, a mis au point des appareils comportant des bouteilles d'air comprimé, que les plongeurs portent sur le dos. Ils exhalent l'air dans des récipients contenant des produits chimiques capables d'absorber le gaz carbonique produit par la respiration. Ces appareils ont donné un nouveau départ au sport de la plongée dès la fin de la guerre. Dans les pays de langue anglaise, on les désigne sous le nom de *scuba* (self-contained underwater breathing apparatus), ce qui traduit bien leur nom français de *scaphandres autonomes*.

Grâce à cet équipement, un plongeur entraîné peut descendre jusqu'à soixante mètres, ce qui est encore très peu par rapport à la profondeur totale de l'Océan.

Avec un scaphandre étanche (le premier a été mis au point en 1830 par Augustus Siebe) de type moderne, on peut descendre à une centaine de mètres. Pour descendre plus bas, il faut un scaphandre rigide qui entoure complètement le corps, et qui puisse se déplacer, autrement dit un *sous-marin*.

Le premier sous-marin effectivement capable de rester un temps appréciable sous l'eau sans noyer son équipage date de 1620 : c'était celui de l'inventeur hollandais Cornelis Drebbel. Toutefois, aucun sous-marin ne pouvait devenir réellement utilisable tant qu'on ne disposerait, pour le propulser, que d'une hélice tournée à la main. Or, on ne pouvait pas employer la vapeur, faute de pouvoir brûler du carburant dans l'atmosphère limitée d'un sous-marin fermé. Il fallait donc un moteur électrique, alimenté par une batterie d'accumulateurs.

Le premier sous-marin électrique de ce type est apparu en 1886. Bien sûr, il fallait recharger périodiquement la batterie, mais entre deux recharges l'appareil pouvait parcourir en plongée plus de 120 kilomètres. Au moment où a commencé la Première Guerre mondiale, les principales puissances européennes disposaient toutes de sous-marins équipés à des fins militaires. Mais ces premiers sous-marins étaient fragiles, et ne pouvaient pas plonger très profondément.

En 1934, Charles William Beebe est parvenu à plonger à près de mille mètres dans sa *bathysphère*, petite cabine aux murs épais, équipée d'oxygène et d'absorbeurs de gaz carbonique.

Mais la bathysphère était un objet inerte, suspendu par un câble à un vaisseau de surface (câble dont la rupture pouvait être fatale). Ce qu'il fallait était un engin indépendant, que l'on puisse manœuvrer à grande profondeur. Le premier fut le *bathyscaphe*, inventé en 1947 par le physicien suisse Auguste Piccard. Sa cabine était construite pour résister à des pressions énormes, et munie d'un lest important de grenaille de fer (automatiquement relâché en cas d'urgence). Cette cabine était soutenue par un « ballon » contenant de l'essence – plus légère que l'eau. Lors de son premier test au large de Dakar, en 1948, le bathyscaphe est descendu sans équipage à 1 500 mètres de profondeur. La même année, le collègue de Beebe, Otis Barton, est descendu à la même profondeur, avec une version améliorée de la bathysphère, appelée *benthoscope*.

Puis, Piccard et son fils Jacques ont construit une version plus perfectionnée du bathyscaphe, qu'ils ont baptisée *Trieste* parce que la ville de Trieste participa au financement. En 1953, Piccard plongea à 3 150 mètres dans la Méditerranée, battant provisoirement le record établi quelques semaines plus tôt par un bathyscaphe de la Marine française piloté par Houot et Willm.

Les États-Unis ont acheté le *Trieste*, et le 14 janvier 1960 Jacques Piccard et Don Walsh (de la Marine américaine) l'ont fait descendre au fond de la fosse des Mariannes, à 11 000 mètres de profondeur, au point le plus bas des fosses abyssales. Là, sous une pression de 1 100 atmosphères, ils ont trouvé des courants et des êtres vivants. Le premier qu'ils ont aperçu était même un vertébré, un poisson voisin du turbot, qui possède des yeux.

En 1964, le bathyscaphe français *Archimède* a plongé dix fois dans la fosse de Porto Rico, la plus profonde de l'Atlantique (8 500 mètres). Là aussi, chaque mètre carré du fond a des habitants. Chose curieuse, les bords de la fosse ne plongent pas progressivement, mais semblent descendre en terrasses, comme les marches d'un gigantesque escalier.

Les calottes glaciaires

Les « bouts du monde » ont toujours fasciné les hommes, et l'un des chapitres les plus aventureux de l'histoire des sciences est celui de l'exploration des régions polaires. Ces régions sont romantiques, spectaculaires, et lourdes d'importance pour le destin de l'humanité, avec leurs étranges aurores boréales, leur froid extrême, et surtout les immenses calottes glaciaires qui jouent un rôle décisif dans nos climats et notre mode de vie.

LE PÔLE NORD

La conquête des pôles est assez tardive dans l'histoire de l'humanité. Elle a commencé au temps des Grandes Découvertes, à l'époque de Christophe Colomb. Les premiers explorateurs de l'Arctique cherchaient surtout le passage du Nord-Ouest, qui permettrait aux navires de contourner l'Amérique. A la poursuite de ce mirage, le navigateur anglais Henry Hudson (au service de la Hollande) explora la baie qui porte son nom et y trouva la mort en 1610. Six ans plus tard, un autre navigateur anglais, William Baffin, donna lui aussi son nom à un bras de mer (voir figure 4.6) qui lui permit de s'approcher à 1 300 kilomètres du pôle. Finalement, entre 1846 et 1848, l'explorateur anglais John Franklin réussit, au prix de sa vie, à contourner le continent : mais le passage du Nord-Ouest enfin découvert n'était guère praticable.

Le demi-siècle suivant allait connaître une série de tentatives pour s'approcher du pôle, par goût de l'aventure et désir de l'atteindre le premier. En 1873, les explorateurs autrichiens Julius Payer et Carl Weyprecht s'en approchèrent à 1 000 kilomètres, découvrant au passage un archipel qu'il nommèrent terre de François-Joseph, en hommage à l'empereur d'Autriche. En 1896, l'explorateur norvégien Fridtjof Nansen s'approcha à moins de 500 kilomètres dans son bateau, qui dériva avec les glaces de la banquise, et il parvint ensuite à parcourir encore cent kilomètres à pied vers le nord avant de devoir faire demi-tour. Enfin, le 6 avril 1909, l'explorateur américain Robert Edwin Peary arriva au pôle.

De nos jours, le pôle a perdu beaucoup de son mystère. On l'a exploré sur la glace, par la voie des airs (premier survol par Richard Evelyn Byrd et Floyd Bennett en 1926) et sous l'eau (plusieurs sous-marins sont passés sous la banquise).

Entre-temps, la grande calotte glaciaire du Groenland a attiré de nombreuses expéditions scientifiques. Wegener est mort au cours de l'une d'elles en novembre 1930. Finalement, on a constaté que la calotte glaciaire avait une surface de un million de kilomètres carrés (plus des trois quarts de la surface totale du Groenland) et que son épaisseur dépassait par endroits 1 500 mètres.

A mesure qu'elle s'accumule, la glace descend vers la mer, où elle donne naissance aux icebergs. Des quelque 16 000 icebergs qui apparaissent chaque année dans l'Arctique, 90 % se sont détachés des bords de la calotte glaciaire du Groenland. Ils dérivent lentement vers le sud, surtout le long de la côte ouest de l'Atlantique. Environ 400 d'entre eux chaque année descendent plus bas que Terre-Neuve et menacent les lignes maritimes.

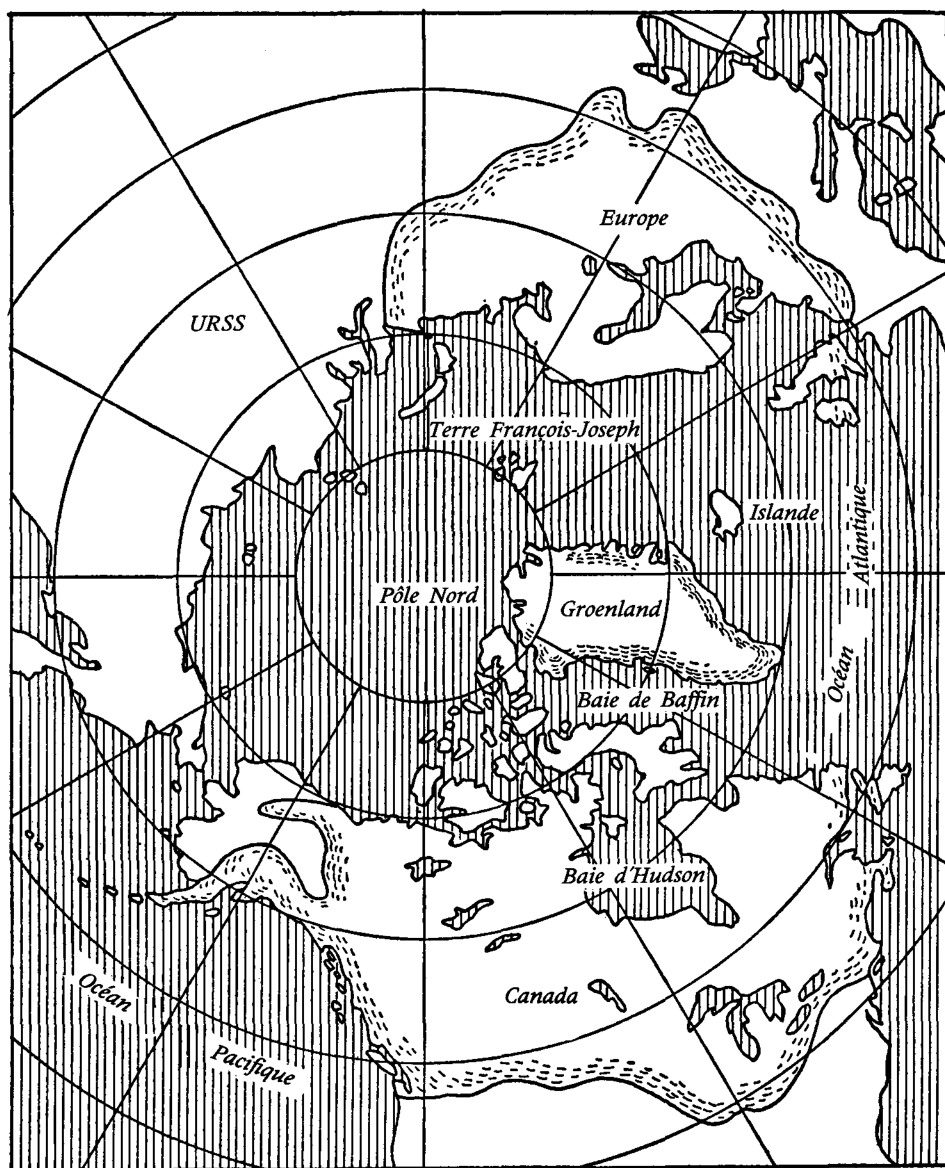


Figure 4.6. Carte des régions polaires arctiques.

Entre 1870 et 1890, des collisions avec des icebergs ont coulé quatorze navires et en ont endommagé quarante autres.

La collision la plus tragique s'est produite plus tard en 1912, lors du voyage inaugural du grand paquebot *Titanic* : heurtant un iceberg, le navire a sombré en entraînant un grand nombre de passagers. Depuis lors, une

coopération internationale s'est mise en place pour surveiller ces monstres inertes. Depuis que fonctionne cette « patrouille des glaces », aucun navire n'a été coulé par un iceberg.

LE PÔLE SUD ET L'ANTARCTIQUE

La calotte glaciaire du pôle Sud est considérablement plus importante que celle du pôle Nord. Elle a une surface sept fois plus grande, et son épaisseur moyenne est de 2 400 mètres, avec des maxima de l'ordre de 5 000 mètres par endroits. Ceci est dû à la grande taille du continent antarctique, environ 13 millions de kilomètres carrés, bien qu'on ne sache pas au juste où s'arrête la terre et où commence la banquise (figure 4.7). Quelques explorateurs soutiennent que la partie occidentale de l'Antarctique, au moins, est un archipel cimenté par la glace, mais pour l'instant, la plupart des spécialistes pensent qu'il s'agit bien d'un continent.

Le célèbre explorateur anglais James Cook fut le premier Européen à traverser le cercle polaire antarctique. En 1773, il fit le tour du continent (c'est peut-être ce voyage qui a inspiré le poème de Samuel Taylor Coleridge, *le Vieux Marin*, publié en 1798, qui décrit un voyage de l'Atlantique au Pacifique en passant par les régions glacées de l'Antarctique).

En 1819, l'explorateur anglais Williams Smith découvrit les Shetland du Sud, à 80 kilomètres de la côte de l'Antarctique. En 1821 une expédition russe, dirigée par Fabian Gottlieb Bellingshausen, découvrit une petite île (île de Pierre Ier) à l'intérieur du cercle antarctique. La même année, l'Anglais George Powell et l'Américain Nathaniel B. Palmer aperçurent pour la première fois une péninsule du continent antarctique proprement dit, qui s'appelle maintenant péninsule de Palmer ou péninsule Antarctique.

Dans les années qui suivirent, les explorateurs se rapprochèrent lentement du pôle. En 1840, l'officier de marine américain Charles Wilkes annonça que tous les points repérés jusque-là devaient faire partie d'une seule masse continentale, et l'avenir montra qu'il avait deviné juste. L'Anglais James Weddell pénétra dans une baie profonde à l'est de la péninsule de Palmer. Cette baie (appelée mer de Weddell) lui permit de descendre jusqu'à 1 500 kilomètres du pôle. Un autre explorateur anglais, James Clark Ross, découvrit l'autre baie profonde de l'Antarctique (la mer de Ross) et parvint ainsi à 1 000 kilomètres du pôle.

De 1902 à 1904, un autre Anglais, Robert Falcon Scott, explora la plate-forme de Ross, région de banquise plus vaste que le Texas, et parvint à 800 kilomètres du pôle. Et en 1909, un autre Anglais encore, Ernest Shackleton, avança sur la glace à 150 kilomètres du but.

Celui-ci allait être finalement atteint le 16 décembre 1911 par l'explorateur norvégien Roald Amundsen. En même temps, Scott fit une nouvelle tentative, désastreuse : il parvint bien au pôle – pour y trouver le drapeau planté trois semaines plus tôt par Amundsen – et mourut de faim et d'épuisement avec ses compagnons sur le chemin du retour.

Vers la fin des années vingt, l'avion rendit possibles des explorations plus systématiques. L'explorateur australien George Hubert Wilkins survola la côte sur 2 000 kilomètres, et en 1929 Richard Evelyn Byrd survola le pôle Sud. A ce moment-là, la première base américaine, Little America I, était déjà installée sur le continent.

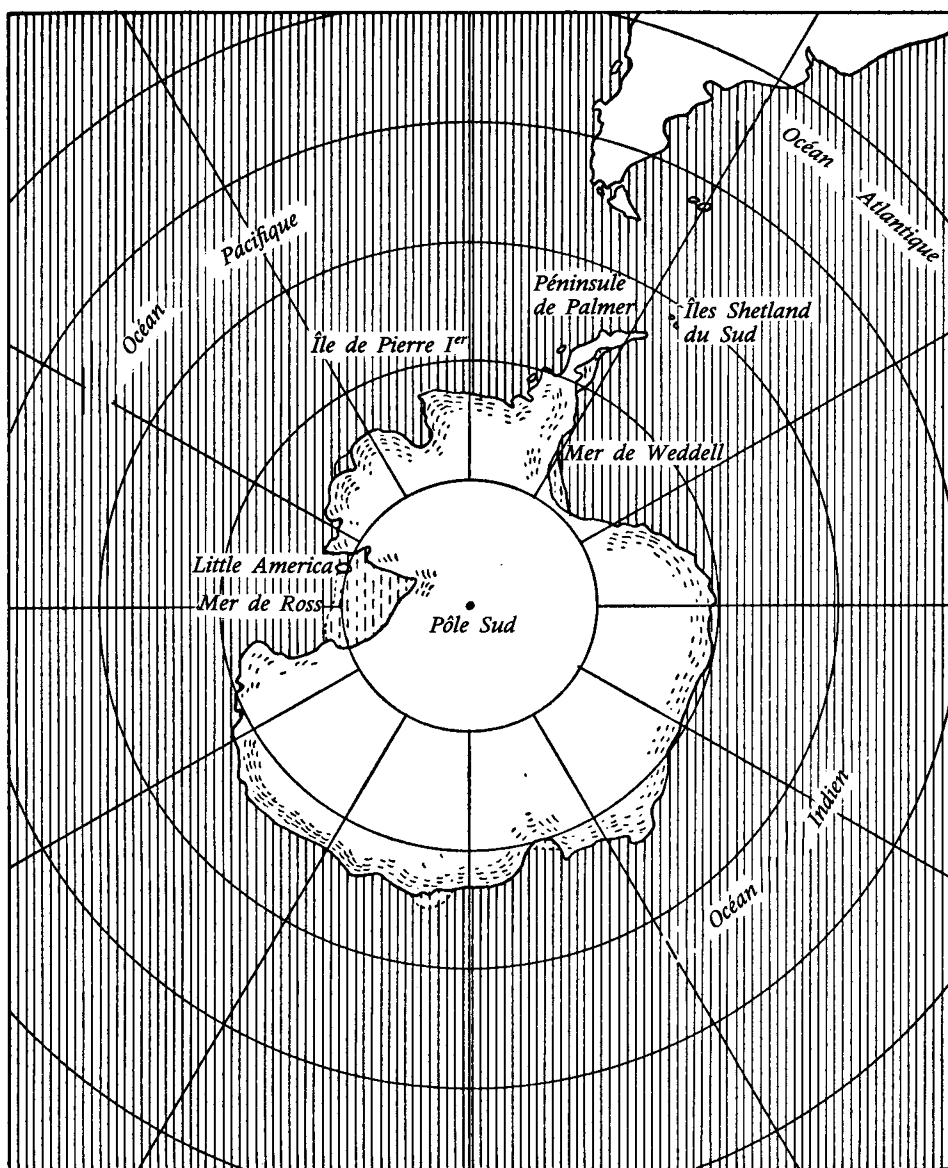


Figure 4.7. Carte des régions polaires antarctiques. L'Antarctique et le Groenland portent l'essentiel des glaciers continentaux actuels. Au plus fort de la dernière glaciation, la glace recouvrait la plus grande partie de l'Europe du Nord et l'Amérique du Nord jusqu'au sud des Grands Lacs.

L'ANNÉE GÉOPHYSIQUE INTERNATIONALE

Les régions polaires nord et sud allaient faire l'objet du programme scientifique international le plus ambitieux des temps modernes. En 1882-83 déjà, plusieurs pays avaient participé à l'Année polaire internationale d'exploration et de recherche scientifique sur des phénomènes comme les aurores boréales et le magnétisme terrestre. Le succès fut tel qu'en 1932-33, l'expérience fut renouvelée avec une seconde Année polaire internationale. En 1950, le géophysicien américain Lloyd Berkner (qui avait fait partie de la première expédition antarctique de Byrd) lança l'idée d'une troisième année polaire. La proposition fut soutenue avec enthousiasme par le Conseil scientifique international. Cette fois, les scientifiques disposaient d'instruments puissants et se posaient une foule de questions nouvelles – sur les rayons cosmiques , la haute atmosphère, les profondeurs océaniques, et même l'éventualité de voyages dans l'espace. On décida donc d'organiser une Année géophysique internationale, du 1er juillet 1957 au 31 décembre 1958 (une période d'activité maximale pour les taches solaires). Cette entreprise rencontra partout un accueil chaleureux : même les États-Unis et l'Union Soviétique, en pleine guerre froide, parvinrent, au nom de la science, à enterrer la hache de guerre.

Si la réalisation la plus spectaculaire, pour le public, de l'Année géophysique internationale a été à coup sûr le lancement de satellites artificiels par l'Union Soviétique, puis par les États-Unis, la récolte de résultats scientifiques ne s'est pas limitée là. En particulier, des équipes du monde entier ont exploré l'Antarctique (à eux seuls, les États-Unis ont établi sept bases sur le continent), sondant les profondeurs de la glace, et en rapportant des échantillons de l'air piégé à des kilomètres de la surface depuis des millions d'années, ainsi que des traces de bactéries. Certaines bactéries, trouvées à trente mètres de profondeur dans la glace, et gelées depuis cent ans, se sont remises à vivre et à se développer normalement. En janvier 1958, l'équipe soviétique a établi une base au pôle d'inaccessibilité (le point situé le plus loin des côtes), à 900 kilomètres du pôle Sud. On y a enregistré des températures records : en août 1960 (au plus fort de l'hiver antarctique), la température est descendue jusqu'à –88 °C, bien au-dessous du point de solidification du gaz carbonique. Pendant les années qui ont suivi, des dizaines de stations permanentes ont continué à travailler dans l'Antarctique.

Parmi toutes les réalisations, la plus spectaculaire a été celle de l'équipe britannique d'exploration dirigée par Vivian Ernest Fuchs et Edmund Percival Hillary, qui a traversé le continent pour la première fois (en disposant, il est vrai, des ressources de la science et de la technique modernes). (Hillary avait été aussi le premier, en 1953, à atteindre, avec le sherpa Tensing, le sommet de l'Everest, le point le plus haut de la Terre.)

Le succès de l'Année géophysique internationale et les sentiments chaleureux entretenus par cette coopération au milieu de la guerre froide débouchèrent en 1959 sur un accord entre douze pays, excluant de l'Antarctique toute activité militaire (y compris les essais de bombes atomiques et le dépôt de déchets radioactifs). Ainsi, l'Antarctique est réservée aux activités scientifiques.

LES GLACIERS

La glace présente sur notre globe, environ 36 millions de kilomètres cubes, couvre 10 % de la surface des terres. De cette glace, 86 % sont entassés dans la calotte antarctique, et 10 % dans celle du Groenland. Les 4 % qui restent forment les glaciers d'Islande, d'Alaska, d'Himalaya, des Alpes, et de quelques autres endroits.

Il y a longtemps que l'on étudie les glaciers alpins. Dès 1820, deux géologues suisses, Ignatz Venetz et Johann von Charpentier, ont remarqué qu'on trouvait des rochers caractéristiques des Alpes centrales éparpillés sur les plaines qui s'étendent au nord. Comment y étaient-ils venus ? Les deux géologues émirent l'hypothèse que les glaciers avaient pu occuper autrefois une surface beaucoup plus considérable, laissant dans leur retraite des rochers et des entassements de débris.

Un zoologiste suisse, Jean Louis Rodolphe Agassiz, se pencha sur la question. Il planta des alignements de piquets dans les glaciers, pour voir s'ils bougeaient. En 1840, il put affirmer sans hésitation qu'ils « coulaient » comme des rivières extrêmement lentes, à raison de quelques dizaines de mètres par an. Entre-temps, ses voyages en Europe lui permirent de trouver des traces de glaciers en France et en Angleterre. Il y rencontra des rochers étrangers à la région où ils se trouvaient, et des rayures sur la pierre qui n'avaient pu être faites que par le frottement des glaciers entraînant des pierres incrustées dans leur surface inférieure.

Agassiz partit pour les États-Unis en 1846 et devint professeur à Harvard. Il découvrit des traces de glaciation en Nouvelle-Angleterre et dans le Middle-West. Dès 1850, il était évident qu'une grande partie de l'hémisphère nord avait été autrefois recouverte par un immense glacier continental. Depuis l'époque d'Agassiz, on a étudié avec soin les dépôts laissés par ce glacier, et on a constaté qu'il avait dû avancer et reculer plusieurs fois durant le dernier million d'années (période que l'on appelle *Pleistocène*).

Les géologues appellent *glaciations pleistocènes* ce que le public connaît sous le nom de « périodes glaciaires ». Des périodes glaciaires, il y en a eu bien avant le Pleistocène, une il y a environ 250 millions d'années, une autre il y a 600 millions d'années, et peut-être une entre les deux, il y a 400 millions d'années. On ne sait pas grand-chose de ces périodes glaciaires anciennes, car leurs traces se sont presque effacées au fil des millions d'années. En tout cas, les périodes glaciaires sont assez rares : moins de 1 % de l'histoire de la Terre.

En ce qui concerne les glaciations du Pléistocène, elles ne semblent pas avoir affecté sensiblement la calotte glaciaire antarctique, qui est pourtant la plus importante de notre temps. C'est qu'elle ne peut s'agrandir que sur la mer, où elle se brise en icebergs. Ceux-ci peuvent devenir plus nombreux et abaisser la température générale des océans. Mais les terres émergées de l'hémisphère sud sont trop loin de l'Antarctique pour que s'y développent des calottes glaciaires (sauf peut-être dans l'extrême-Sud des Andes).

Il en va tout autrement de l'hémisphère nord. Là, des continents se resserrent autour du pôle. C'est donc là que l'extension des glaciers est la plus spectaculaire. Aussi, on ne parle des glaciations pléistocènes qu'en ce qui concerne l'hémisphère nord. A cette époque-là, en plus de la calotte glaciaire du Groenland, qui existe toujours, il y en avait trois autres,

chacune couvrant près de trois millions de kilomètres carrés, respectivement au Canada, en Scandinavie et en Sibérie.

Peut-être à cause de sa proximité du Groenland (germe des glaciations nordiques) le Canada est-il plus affecté que la Scandinavie et surtout que la Sibérie. Croissant à partir du nord-est, la calotte glaciaire du Canada laisse à découvert l'Alaska et la côte pacifique de l'Amérique du Nord, mais couvre la plus grande partie du Nord des États-Unis. A son extension maximale, la limite des glaces va de Seattle (État de Washington) jusqu'à l'île de Long Island (en face de New York). Rappelons que New York est à la latitude de Naples, et que les glaciations européennes ne sont jamais allées jusque-là.

L'un dans l'autre, à leur plus grande extension, les glaciers couvrent à peu près 30 % des terres émergées, soit trois fois leur surface actuelle.

En examinant soigneusement les sédiments déposés par les glaciers, on constate qu'ils ont avancé et reculé quatre fois. Chacune des quatre périodes a duré de 50 000 à 100 000 ans, les périodes chaudes *interglaciaires* étant nettement plus longues.

La quatrième et dernière des glaciations a atteint son maximum il y a environ 18 000 ans. Puis, un lent repli s'ést amorcé. « Lent », en l'occurence, veut dire que la glace, aux endroits les plus favorisés, reculait à peu près de 80 mètres par an – mais il lui arrivait encore d'avancer de temps en temps.

Il y a une dizaine de milliers d'années, à l'époque où la civilisation commençait à s'esquisser au Proche-Orient, les glaciers nordiques ont commencé leur retraite finale. Deux mille ans plus tard les Grands Lacs étaient à découvert, et finalement, il y a environ cinq mille ans (au moment où apparaissait l'écriture), la frontière des glaces avait pratiquement atteint sa position actuelle.

Les allées et venues des glaciers coïncident bien sûr avec des variations climatiques à l'échelle de la planète, qui laissent des traces perceptibles, mais de plus, ces allées et venues se traduisent par des changements de forme des continents. Par exemple, si les calottes glaciaires actuelles du Groenland et de l'Antarctique fondaient, le niveau des océans s'élèverait de 70 mètres. Ceci inonderait la plupart des grandes villes de tous les continents : l'eau monterait jusqu'au vingtième étage des gratte-ciel de Manhattan. En revanche, l'Alaska, le Canada, la Sibérie, le Groenland et même l'Antarctique deviendraient plus habitables.

Au plus fort d'une glaciation, c'est l'inverse qui se produit. Il y a tellement d'eau piégée dans les calottes glaciaires continentales (trois ou quatre fois plus qu'il y en a actuellement) que le niveau de la mer est au moins à 140 mètres en dessous de ce qu'il est de nos jours. Aussi les plateaux continentaux sont-ils découverts.

Les plateaux continentaux sont des portions peu profondes des océans, au voisinage des continents. En effet, le fond de la mer descend doucement jusqu'à une profondeur de l'ordre de 130 mètres. Ensuite, la pente devient beaucoup plus abrupte, et on atteint très vite des profondeurs importantes. Du point de vue de la structure, les plateaux continentaux font partie des continents : la vraie limite de ceux-ci est le bord du plateau continental. Actuellement, il y a assez d'eau dans les océans pour recouvrir ces limites.

Et la surface en question est loin d'être négligeable. La largeur du plateau continental est très variable : elle est beaucoup plus grande sur la côte est des États-Unis que sur la côte ouest, qui est à la limite d'une plaque

tectonique. En moyenne, le plateau continental a une largeur de l'ordre de 80 kilomètres, ce qui lui donne une surface totale de 25 millions de kilomètres carrés. Autrement dit, une surface continentale équivalente à celle de l'Union Soviétique est actuellement en dessous du niveau de la mer.

C'est cette surface qui est à découvert au maximum des périodes de glaciation, par exemple au maximum des dernières glaciations. On a tiré du plateau continental, à des kilomètres du rivage et à des mètres de profondeur, des fossiles d'animaux terrestres, comme des dents d'éléphant.

D'autre part, si le Nord des continents est couvert de glace, il pleut davantage dans le Sud : lors de la dernière glaciation, le Sahara était une savane couverte d'herbe. Son assèchement, consécutif au retrait des calottes glaciaires, n'est pas bien antérieur au début des temps historiques.

Ainsi, en matière d'habitabilité, il y a une sorte de compensation. Quand le niveau de la mer baisse, des régions continentales importantes deviennent des déserts de glace, mais les plateaux continentaux deviennent habitables, ainsi que les déserts actuels. Quand le niveau de la mer s'élève, les terres basses sont inondées, mais les régions nordiques deviennent habitables : c'est maintenant là que le désert recule.

C'est pourquoi les glaciations n'étaient pas forcément des périodes de désolation et de catastrophes. Toute la glace de toutes les calottes au maximum de la glaciation représente seulement 0,35 % de l'eau présente dans les océans. Aussi les glaciations ont-elles peu d'effet sur l'Océan. Bien sûr, les régions peu profondes sont très réduites, et ce sont des régions très peuplées. Mais par ailleurs, les eaux tropicales sont plus fraîches (de 2 à 5 degrés) qu'elles le sont actuellement, c'est-à-dire qu'elles dissolvent davantage d'oxygène, et permettent une vie plus riche.

D'autre part les allées et venues de la glace sont très lentes, et cette lenteur donne le temps aux espèces animales d'émigrer peu à peu vers le sud ou vers le nord, ou même de s'adapter aux changements de climat : les glaciations étaient les périodes florissantes du mammouth laineux...

Enfin, les variations ne sont pas aussi démesurées qu'on pourrait le penser, parce que la glace ne fond jamais complètement. La calotte antarctique existe, sans changement important, depuis une vingtaine de millions d'années, et contribue à limiter les fluctuations du niveau de la mer et de la température.

Cela ne veut pas dire que l'avenir ne nous donne aucune inquiétude. Il n'y a aucune raison d'exclure l'éventualité d'une cinquième glaciation, avec les problèmes qu'elle poserait. Lors de la dernière, les hommes étaient peu nombreux. C'étaient des chasseurs, qui pouvaient facilement se déplacer vers le nord ou vers le sud, en suivant leur gibier. Lors de la prochaine, les hommes seraient à coup sûr (ils le sont déjà) plus nombreux, et relativement sédentarisés par leurs villes et autres constructions. De plus, il est possible que certaines manifestations du progrès technologique influent sur l'avance ou le recul des glaciers.

LES CAUSES DES GLACIATIONS

C'est la question la plus importante que pose le phénomène. Qu'est-ce qui fait avancer et reculer la glace, et pourquoi les périodes glaciaires sont-elles relativement brèves (la dernière occupant seulement l'un des cent derniers millions d'années) ?

Il suffit d'un petit changement de température pour provoquer ou terminer une glaciation – juste ce qu'il faut pour qu'il tombe un peu plus de neige en hiver qu'il n'en fond en été, ou l'inverse. On estime qu'une baisse de température moyenne annuelle de 3,5 °C suffirait à mettre les glaciers en marche, tandis qu'une augmentation équivalente découvrirait en quelques siècles les rochers du Groenland et de l'Antarctique.

En effet, il se produit – si l'on ose dire – un phénomène de boule de neige. Une chute de température suffisante pour augmenter en quelques années la surface des glaces entraîne une chute de température plus rapide. En effet, la glace reflète la lumière du Soleil beaucoup mieux que le rocher ou le sol nu : 90 %, alors que le sol nu en renvoie moins de 10 %. Si la surface de glace augmente légèrement, la réflexion des rayons solaires augmente, leur absorption diminue, donc la température baisse, et la glace s'étend plus vite.

Inversement, si la température augmente légèrement – assez pour faire reculer un peu la glace – la lumière du Soleil est moins réfléchie, donc plus absorbée, augmentant la température et la vitesse du repli des glaces.

Mais, dans les deux cas, qu'est-ce qui déclenche le processus ?

L'une des réponses possibles est fondée sur le fait que l'orbite terrestre n'est pas tout à fait stable, et ne se reproduit pas d'une façon immuable au fil des années. Par exemple, l'époque du périhélie change. En ce moment, le périhélie (moment où la Terre est la plus proche du Soleil) se place peu après le solstice d'hiver. Mais le périhélie se décale chaque année, et fait le tour de l'orbite en 21 310 ans. D'autre part, la direction de l'axe des pôles change, et décrit un cône dans le ciel (c'est la précession des équinoxes) en 25 780 ans. Enfin, l'inclinaison de l'axe varie aussi un peu, oscillant légèrement autour de sa valeur moyenne.

Tous ces changements influent un peu sur la température de la Terre – très peu, mais peut-être assez pour déclencher de temps en temps l'avance des glaciers ou leur retrait.

En 1920, un physicien yougoslave, Milutin Milankovich, a avancé l'idée d'un cycle de ce type, durant 40 000 ans, avec un « Grand Printemps », un « Grand été », un « Grand Automne » et un « Grand Hiver » durant chacun 10 000 ans. Selon cette théorie, la Terre serait particulièrement susceptible d'entrer en glaciation pendant le « Grand Hiver », et le processus se déclencherait à ce moment, pour peu que d'autres facteurs y soient aussi favorables. Une fois glacée, la Terre se « dégèlerait » le plus facilement lors du « Grand été », à condition que d'autres facteurs s'y prêtent.

La proposition de Milankovich n'a pas soulevé l'enthousiasme au moment de son apparition. Mais en 1976, le problème a attiré l'attention de J.D. Hays et John Imbrie (États-Unis) et N.J. Shackleton (Grande-Bretagne). Ils disposaient de longues « carottes » sédimentaires rapportées de deux endroits différents de l'océan Indien – des bas-fonds éloignés de la côte, qui ont peu de chances d'avoir été contaminés par des matériaux descendus des zones côtières, ou même de zones voisines plus hautes.

Ces carottes étaient constituées par des matériaux déposés régulièrement pendant 450 000 ans. Bien sûr, en bas de la carotte on a trouvé les sédiments les plus anciens. On a pu y étudier les squelettes de petits animaux unicellulaires, dont les différentes espèces prolifèrent à différentes températures : à partir de la nature des squelettes observés à différents niveaux, on a pu estimer la température aux époques correspondantes.

D'autre part, les atomes d'oxygène se présentent sous deux variétés principales, dont les proportions dépendent de la température. En

mesurant ces proportions à différents niveaux de la carotte, on a pu estimer la température aux époques correspondantes.

Or, les deux estimations étaient en parfait accord, et semblaient indiquer quelque chose qui ressemblait beaucoup au cycle de Milankovich. Il est donc possible que la Terre connaisse de temps en temps un « Grand Hiver » glacé, de même qu'elle a chaque année un hiver neigeux.

Seulement, pourquoi le cycle de Milankovich aurait-il fonctionné pendant le Pléistocène, et pas pendant les deux cents millions d'années antérieurs, au cours desquels on n'a décelé aucune glaciation ?

En 1953, Maurice Ewing et William L. Donn ont suggéré que cela pourrait être dû à la géographie de l'hémisphère nord. La région arctique est un océan enserré de toutes parts par des masses continentales.

Supposons cet océan un peu moins froid qu'aujourd'hui, et offrant une étendue liquide au lieu de la banquise. Il formerait alors une source de vapeur d'eau, qui, une fois refroidie dans la haute atmosphère, retomberait sous forme de neige. La neige qui tomberait dans l'océan y fondrait, mais celle qui tomberait sur les continents voisins s'accumulerait, déclenchant la glaciation. L'océan Arctique gèlerait.

La glace ne libère pas autant de vapeur que l'eau liquide à la même température (nous citons toujours la théorie de Maurice Ewing et William L. Donn). Une fois l'océan gelé, il y a moins de vapeur dans l'air, et moins de chutes de neige. Les glaciers commencent leur retraite, et si cela déclenche la déglaciation, alors ce retrait s'accélère.

Ainsi, on peut imaginer que le cycle de Milankovich ne déclenche des périodes de glaciation que lorsqu'il y a un océan enserré dans les terres à l'un ou à l'autre pôle (ou aux deux).

Il pourrait ainsi s'écouler des centaines de millions d'années sans qu'existe un tel océan « fermé », puis le mouvement des plaques tectoniques (quelques dizaines de kilomètres par million d'années) créerait brusquement une situation convenable, pendant un million d'années, période durant laquelle les glaciers avanceraient et reculeraient régulièrement. Cette suggestion intéressante n'est pas encore universellement adoptée.

Il y a évidemment des changements moins réguliers dans la température de la Terre, et des causes plus irrégulières de réchauffement et de refroidissement. Le chimiste américain Jacob Bigeleisen, collaborateur de H.C. Urey, a mesuré les proportions des deux variétés d'oxygène dans les fossiles d'animaux marins, de manière à mesurer la température de l'eau dans laquelle vivaient ces animaux. Vers 1950, Urey et son équipe ont perfectionné cette technique à un point tel qu'en analysant les couches de la coquille d'un fossile de plusieurs millions d'années (une sorte de calmar disparu depuis), ils ont pu déterminer que l'animal était né en été, avait vécu quatre ans, et était mort au printemps.

Ce « thermomètre » a permis d'établir qu'il y a cent millions d'années, la température moyenne de l'Océan était voisine de 21 °C. Dix millions d'années plus tard, elle était descendue progressivement jusqu'à 16 ° pour remonter à 21° pendant les dix millions d'années suivants. Depuis, cette température baisse constamment. Quelle que soit la cause de ce refroidissement, c'est peut-être aussi un facteur de la disparition des dinosaures (sans doute adaptés à des climats tempérés et peu variables), et de l'expansion des mammifères et des oiseaux qui, animaux à sang chaud, peuvent maintenir constante leur température interne.

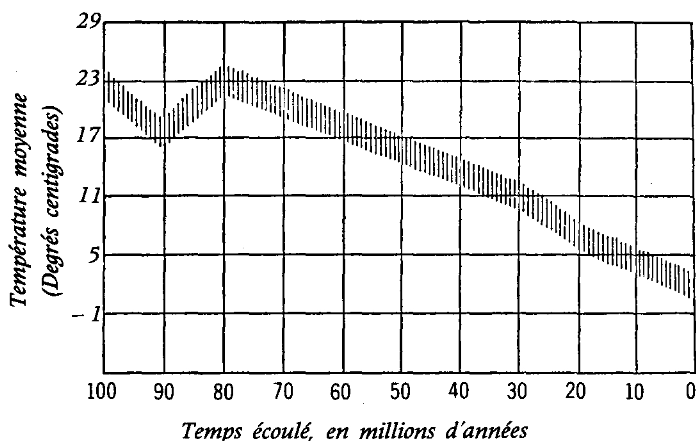


Figure 4.8. Les températures océaniques pendant les cent derniers millions d'années.

Cesare Emiliani, utilisant la technique d'Urey, a étudié les coquilles des foraminifères rapportées dans des carottes extraites du plancher océanique. Il a trouvé que la température moyenne de l'Océan était de l'ordre de 15,5 °C il y a trente millions d'années, de 12 °C il y a vingt millions d'années, et de 2 ° actuellement (figure 4.8).

Quelle est la cause de ces lents changements de température ? Une explication possible se fonde sur l'« effet de serre » du gaz carbonique. Celui-ci absorbe assez fort l'infrarouge. Donc, quand il y a beaucoup de gaz carbonique dans l'atmosphère, cela tend à empêcher la chaleur accumulée par la terre pendant la journée de s'échapper pendant la nuit. Dans ce cas, la chaleur s'accumule. Au contraire, quand la teneur de l'atmosphère en gaz carbonique diminue, la terre se refroidit régulièrement.

Si la teneur actuelle en gaz carbonique doublait (passant de 0,03 % à 0,06 % de la masse de l'atmosphère), ce petit changement suffirait à élever de 3 ° la température moyenne de la Terre, et à fondre complètement et rapidement les calottes glaciaires continentales. Si au contraire la concentration en gaz carbonique diminuait de moitié, la température baisserait suffisamment pour que les glaciers arrivent à nouveau dans la région de New York.

Les volcans libèrent dans l'atmosphère de grandes quantités de gaz carbonique. En revanche, les roches exposées à l'air absorbent du gaz carbonique, ce qui donne du calcaire. Voilà donc, peut-être, deux mécanismes capables d'expliquer les changements à long terme des climats. Une période d'activité volcanique supérieure à la moyenne pourrait augmenter la concentration en gaz carbonique de l'atmosphère, et déclencher un échauffement de la Terre. Au contraire, une période de formation de montagnes exposerait de grandes surfaces de roche « neuve » à l'atmosphère, et diminuerait la concentration du gaz carbonique. Ce dernier processus a pu se produire à la fin du *Mésozoïque* (l'époque des reptiles), il y a quelque quatre-vingts millions d'années, au moment où a commencé la lente diminution de la température terrestre.

Mais quelle que soit la cause des glaciations passées, il semble bien que les hommes puissent modifier le climat dans la période qui commence. Le physicien américain Gilbert N. Plass pense que les glaciations sont désormais impossibles, parce que les cheminées de notre civilisation produisent une quantité considérable de gaz carbonique. Cent millions de cheminées, en effet, déversent constamment du gaz carbonique dans l'atmosphère : en tout, cela fait six milliards de tonnes par an, deux cents fois ce que produisent les volcans. D'après Plass, cela a augmenté de 10 % la quantité de gaz carbonique présente dans l'atmosphère depuis 1900, et devrait aboutir à une augmentation équivalente d'ici à l'an 2 000. Cet accroissement de l'« effet de serre » devrait, selon lui, élever la température moyenne de la Terre de 1,1 % par siècle. Effectivement, c'est bien le rythme d'augmentation de la température pendant la première moitié du XXe siècle, selon les résultats dont nous disposons, et qui concernent surtout l'Europe et l'Amérique du Nord. Si le réchauffement continue au même rythme, les glaciers continentaux pourraient bien disparaître en un siècle ou deux.

D'ailleurs, les recherches menées pendant l'Année géophysique internationale semblent bien indiquer que les glaciers reculent un peu partout. En 1959, l'un des grands glaciers de l'Himalaya avait reculé de 200 mètres depuis 1935. d'autres avaient reculé de 300 ou même 600 mètres. Les poissons adaptés aux eaux froides émigrent vers le nord, et les arbres tropicaux avancent dans la même direction. Le niveau de la mer s'élève un peu chaque année, comme on s'attend à ce qu'il le fasse si les glaciers fondent. Il a déjà monté au point de menacer d'inondation le métro de New York quand une tempête fait rage à marée haute.

Et pourtant, il semble bien que la température baisse un peu depuis les années quarante, ayant déjà compensé la moitié de l'augmentation survenue entre 1880 et 1940. C'est peut-être dû à la quantité croissante de poussières et de « smog » présente dans l'atmosphère, qui empêchent la lumière du Soleil d'atteindre la Terre, et lui font, en quelque sorte, de l'ombre. Il semblerait que deux variétés de pollution atmosphérique humaine soient actuellement en « équilibre », au moins sous ce rapport, et au moins momentanément.

Chapitre 5

L'atmosphère

Les couches d'air

Pour Aristote, le monde était formé de quatre couches successives, correspondant aux quatre éléments : la terre (la boule solide), l'eau (l'Océan), l'air (l'atmosphère), et le feu (couche externe invisible, sauf de temps en temps sous forme d'éclairs, lors des orages). Au-delà de ces couches, le monde d'Aristote était formé d'un élément céleste et parfait, qu'il appelait *éther* (en latin, on l'appelait d'un nom qui devait, au Moyen Age, donner la « quintessence », ou cinquième élément).

Il n'y avait pas de place dans ce modèle pour le vide : l'eau commençait à la fin de la terre, l'air à la limite des deux couches précédentes, le feu à la fin de l'air, et à la limite du feu l'éther commençait et s'étendait jusqu'au bout de l'Univers. Selon les Anciens, « la nature a horreur du vide ».

MESURER L'AIR

La pompe destinée à tirer l'eau des puits, une invention très ancienne, semble illustrer admirablement cette horreur du vide (figure 5.1). Un piston coulisse étroitement dans un cylindre. Quand on abaisse le bras de la pompe, le piston monte, laissant un vide au bas du cylindre. Mais la nature ayant horreur de ce vide, l'eau se fraie un passage à travers la soupape du fond et se précipite pour le combler. Les coups de pompe successifs élèvent l'eau de plus en plus haut dans le cylindre, jusqu'au moment où elle se déverse à la sortie de la pompe.

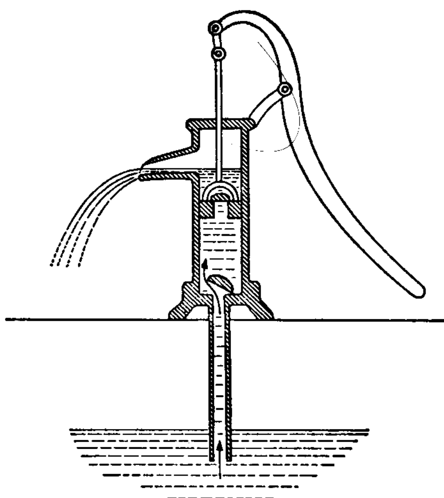


Figure 5.1. Principe de la pompe à eau. Quand on abaisse la poignée, le piston s'élève, créant un vide partiel dans le cylindre, et l'eau y pénètre à travers la soupape. Après plusieurs coups de pompe, le niveau d'eau atteint le tuyau de sortie.

Suivant la théorie aristotélicienne, on pourrait ainsi élever l'eau de n'importe quelle hauteur. Mais les mineurs, qui devaient pomper l'eau du fond des mines, savaient bien qu'ils avaient beau pomper fort et longtemps, ils ne pouvaient jamais la faire monter de plus de dix mètres au-dessus de son niveau naturel.

Galilée se pencha sur le problème à la fin de sa longue et fructueuse carrière. La seule conclusion à laquelle il parvint, c'est qu'apparemment la nature n'avait horreur du vide que jusqu'à un certain point. Il se demanda si la hauteur obtenue serait plus faible avec un liquide plus dense que l'eau, mais il mourut avant d'avoir eu le temps de faire l'expérience.

Cette expérience, deux de ses élèves, Evangelista Torricelli et Vincenzo Viviani, allaient la réaliser en 1644. Choisissant le mercure (de densité 13,6), ils en remplirent un tube d'un mètre ouvert à un bout, le fermèrent, le retournèrent sur une cuve de mercure et le débouchèrent. Le mercure du tube commença à descendre dans la cuve, mais quand sa surface arriva à une hauteur de 76 centimètres au-dessus du niveau de la cuve, le mercure cessa de descendre.

C'était le premier *baromètre.* Les baromètres modernes à mercure en diffèrent très peu. Il ne fallut pas longtemps pour constater que la hauteur du mercure dans le tube n'était pas toujours la même. Torricelli se mit à étudier ces variations, créant ainsi une nouvelle science, la *météorologie.* Bien plus tard, après 1660, l'Anglais Robert Hooke signala que la hauteur de mercure diminuait avant un orage.

Mais qu'est-ce donc qui empêchait le mercure de tomber dans la cuve ? Viviani suggéra que c'était le poids de l'atmosphère, qui appuyait sur le mercure de la cuve. C'était là une idée révolutionnaire, car suivant la théorie aristotélicienne l'air n'avait pas de poids, et rejoignait seulement en toute circonstance la couche qui lui était propre. Et voilà que Viviani

suggérait qu'une colonne d'eau de dix mètres, ou une colonne de mercure de 76 centimètres (qui ont le même poids) étaient équilibrées par une colonne d'air de même section s'étendant jusqu'à la limite supérieure de l'atmosphère.

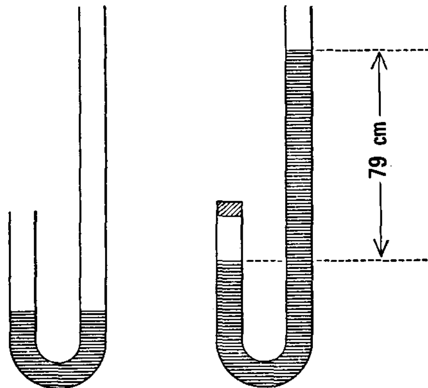
Cette idée souleva une vive controverse dans les milieux scientifiques, controverse qui ne cessa qu'après les démonstrations précises et concluantes d'un jeune mathématicien français, Blaise Pascal. En 1648, celui-ci imagina une expérience, consistant à emporter un tube de Torricelli en haut d'une montagne : soumis à un poids d'air plus faible, le mercure devait y monter moins haut qu'au pied de la montagne. Le beau-frère de Pascal, Florin Périer, réalisa l'expérience au puy de Dôme, avec un succès total.

Mais même avant les travaux décisifs de Pascal, l'expérience de Torricelli avait créé le premier vide artificiel : l'espace qui surmontait le mercure, en haut du tube, était bien vide (à part une quantité infime de vapeur de mercure).

On se mit aussitôt à étudier les propriétés du vide. En 1650, l'Allemand Athanasius Kircher montra que le son ne se transmettait pas à travers le vide. Quelques années plus tard, Robert Boyle prouva que des objets très légers tombaient aussi vite que des objets lourds dans le vide, confirmant ainsi la théorie du mouvement établie par Galilée en opposition aux idées aristotéliciennes.

Ainsi, Pascal a définitivement montré que l'air est pesant. Il a même calculé la masse totale de l'atmosphère. Or, si l'air a un poids bien défini, il semble que l'atmosphère doive avoir une hauteur bien définie. Il y a un kilo d'air au-dessus de chaque centimètre carré de surface terrestre (au niveau de la mer). Par conséquent, si l'air a partout la même densité qu'au voisinage de la mer, l'atmosphère a un peu plus de sept kilomètres d'épaisseur. Mais en 1662, Boyle constata que cette estimation était fausse, car la densité de l'air varie avec la pression. Il utilisa un tube en forme de J, dans lequel un peu d'air était enfermé dans la branche courte par le mercure versé dans l'autre branche. A mesure que l'on ajoutait du mercure, la poche d'air s'amenuisait. En mesurant ces variations, Boyle montra que si la pression exercée sur l'air doublait, son volume diminuait de moitié. Autrement dit, le volume d'une masse donnée d'air varie en raison inverse de sa pression (voir figure 5.2). Cette loi, découverte

Figure 5.2. Expérience de Boyle. En bouchant la branche de gauche et en rajoutant du mercure dans celle de droite, on comprime l'air emprisonné. Le volume de cet air varie de façon inversement proportionnelle à la pression (loi de Boyle-Mariotte).



indépendamment à la même époque par le physicien français Edme Mariotte, s'appelle *loi de Boyle* dans les pays de langue anglaise, et *loi de Mariotte* dans les pays de langue française... Quoi qu'il en soit, c'est l'un des fondements des études ultérieures sur la matière, qui devaient finalement déboucher sur la théorie atomique.

Comment la pression de l'air varie-t-elle avec l'altitude ? En supposant que la température est la même dans toute l'atmosphère, on peut montrer que cette pression est divisée par dix chaque fois que l'on s'élève de vingt kilomètres. Autrement dit, à vingt kilomètres d'altitude, la colonne de mercure dans le tube de Torricelli ne serait plus de 76 centimètres, mais de 7,6 centimètres. A quarante kilomètres, de 7,6 millimètres, et ainsi de suite. A 160 kilomètres, la hauteur de la colonne ne serait plus que de 0,0000076 millimètres. Évidemment, cela semble peu, et pourtant, la masse de l'air situé autour de la Terre à plus de 160 kilomètres d'altitude est encore de six millions de tonnes.

En réalité, tous ces chiffres sont seulement approchés, parce que la température de l'air varie avec l'altitude. Néanmoins, ils aident à se faire une idée des choses, et à se rendre compte que l'atmosphère n'a pas de limite bien définie. Elle se dilue progressivement dans le vide spatial. On a détecté des météores à 160 kilomètres d'altitude, là où la pression est un milliard de fois plus faible qu'au niveau de la mer. Ainsi, cet air extrêmement ténu suffit pourtant à échauffer par frottement ces petits grains jusqu'à l'incandescence. Et les aurores boréales, formées de gaz rendu lumineux par le bombardement de particules venues de l'espace, s'étendent jusqu'à mille kilomètres au-dessus de la Terre.

VOYAGER DANS LES AIRS

Il semble que l'humanité, dès ses débuts, ait rêvé de voyager à travers les airs. De fait, le vent emporte à travers les airs de petits objets légers (plumes, feuilles, graines). D'autre part, certains animaux planent, comme les écureuils volants, les phalangers et les poissons volants, ou même volent vraiment comme les insectes ailés, les chauves-souris et les oiseaux.

Ce désir des hommes laisse des traces dans les mythes et les légendes. Les dieux et les démons se déplacent couramment à travers les airs, les anges et les fées sont souvent munis d'ailes, et puis il y a Icare (qui a donné son nom à un astéroïde, comme nous l'avons vu au chapitre 3), le cheval volant Pégase, et les tapis volants des légendes orientales.

Le premier engin capable de s'élever dans l'air à des hauteurs importantes pendant une durée importante a été le cerf-volant, fait de papier ou d'un matériau léger du même type, tendu sur une carcasse légère de bois, muni d'une queue pour la stabilité et d'un long fil permettant de le tenir. Son invention (si on laisse de côté la légende qui l'attribue au philosophe grec Archytas, au IVe siècle av. J.-C.) remonte au Moyen Age chinois.

Pendant des siècles, on s'est servi des cerfs-volants pour s'amuser, bien qu'on eût pu en concevoir des usages utiles : emporter très haut une lanterne, transporter à travers un ravin une ficelle qui puisse être ensuite utilisée pour tirer une corde...

Mais la première tentative pour utiliser un cerf-volant à des fins scientifiques n'est venue qu'en 1749 : un astronome écossais, Alexander

Wilson, a attaché des thermomètres à des cerfs-volants, dans l'espoir de mesurer les températures à des hauteurs variables. Vers la même époque (en 1752), Benjamin Franklin a tiré beaucoup plus de notoriété de son cerf-volant (nous en parlerons au chapitre 9).

Les cerfs-volants, en Europe au moins, étaient trop petits pour emmener un passager (jusqu'à la fin du XIX[e] siècle), mais cet objectif allait être atteint, par une méthode toute différente, dès le XVIII[e] siècle.

En 1782 les frères Joseph Michel et Jacques Etienne Montgolfier allumèrent un feu sous un sac muni d'une ouverture en bas. Le sac se remplit d'air chaud et s'éleva dans l'air : c'était la première montgolfière. La deuxième emmenait déjà des animaux, et la troisième deux passagers humains, dès l'automne 1783. Quelques semaines plus tard, le physicien français Jacques Alexandre César Charles s'éleva avec un ballon gonflé à l'hydrogène : ce gaz étant quatorze fois plus léger que l'air, chaque kilo d'hydrogène permettait de soulever treize kilos de poids mort.

Dès 1804, le physicien français Louis Joseph Gay-Lussac monta à 7 000 mètres et rapporta des échantillons d'air raréfié. Mais les ballons de l'époque servaient surtout à des acrobaties comme celles du Français Jean-Pierre Blanchard, qui réalisa dès 1785 le premier parachute.

Dans une nacelle ouverte, on ne pouvait pas s'élever beaucoup plus haut que Gay-Lussac. En 1875, trois hommes s'élevèrent à plus de 9 000 mètres, mais un seul d'entre eux, Gaston Tissandier, survécut à la raréfaction de l'air. Il put ainsi décrire les symptômes du manque d'oxygène, écrivant la première page de la *médecine aérienne.* Pour aller plus haut et rapporter des mesures de température et de pression des zones les plus élevées, il fallait des ballons munis d'appareils automatiques, mais sans équipage : ce furent les ballons-sondes, qui entrèrent en service en 1892.

On constata ainsi que pendant les premiers kilomètres de l'ascension, la température tombait, comme on s'y attendait. Vers 11 000 mètres, elle descendait à — 55° C. Mais plus haut, à la surprise générale, elle cessait de diminuer, et même remontait un peu.

Le météorologiste français Léon Philippe Teisserenc de Bort émit en 1902 l'idée que l'atmosphère pouvait avoir deux niveaux : un niveau inférieur turbulent, avec des nuages, des vents, des tempêtes et tous les changements de temps que nous connaissons (en 1908, il baptisa ce niveau *troposphère,* à partir du mot grec signifiant « sphère des changements »), et une couche supérieure tranquille, contenant des sous-couches de gaz légers, hélium et hydrogène (la *stratosphère* ou « sphère des couches »). Quant à la limite entre les deux, l'altitude à laquelle la température cesse de baisser, Teisserenc de Bort l'appela *tropopause* (« fin des changements »). On a constaté depuis que son altitude variait, de 15 000 mètres environ au-dessus du niveau de la mer dans les régions équatoriales à 8 000 mètres dans les régions polaires.

Pendant la Seconde Guerre mondiale, les bombardiers américains volant à haute altitude ont permis de découvrir un phénomène spectaculaire juste au-dessous de la tropopause, le *courant-jet,* formé de vents très forts et très stables soufflant d'ouest en est à des vitesses pouvant atteindre 800 km/h. En réalité, il y a deux courants-jets, l'un dans l'hémisphère nord à la latitude de l'Europe, de la Chine et des États-Unis, l'autre dans l'hémisphère sud à la latitude de l'Argentine et de la Nouvelle-Zélande. Ces courants se déplacent vers le nord ou le sud, provoquant souvent des tourbillons beaucoup plus au nord ou au sud que leur parcours habituel. Les avions

modernes tirent profit de ces courants rapides, mais ce n'est pas là leur seule importance : ils influent énormément sur le mouvement des masses d'air plus basses. Aussi leur connaissance aide-t-elle beaucoup les prédictions météorologiques.

Les hommes ne pouvant pas survivre dans l'atmosphère froide et ténue des hautes altitudes, il a fallu inventer des cabines étanches, dans lesquelles on puisse maintenir une température et une pression correspondant aux conditions à la surface de la Terre. C'est ainsi qu'en 1931 les frères Piccard, Auguste et Jean Félix, le premier devant plus tard inventer le bathyscaphe, se sont élevés à 18 000 mètres dans un ballon supportant une nacelle étanche. Plus tard, des ballons de matière plastique, plus légère et moins poreuse que la soie, ont permis de monter plus haut et d'y rester plus longtemps. En 1938, le ballon *Explorer II* s'est élevé à 21 000 mètres, et depuis des ballons habités sont montés à 38 000 mètres, et des ballons sans équipage à plus de 50 000 mètres d'altitude.

Ces vols ont permis d'étudier la température à des hauteurs plus élevées que précédemment. On a ainsi constaté que la zone à température presque constante ne s'étendait pas indéfiniment vers le haut. Vers 30 000 mètres la stratosphère se termine, et plus haut la température commence à croître !

Cette *haute atmosphère*, située au-dessus de la stratosphère, contient seulement 2 % de la masse totale de l'air. Son exploration n'a commencé que dans les années quarante, grâce à un nouveau type de véhicule, la fusée (voir chapitre 3).

La manière la plus directe de lire les instruments qui enregistrent les conditions à haute altitude consiste à les rapporter au sol pour les regarder. Avec des instruments portés par des cerfs-volants, c'est facile, mais avec les ballons ce l'est déjà moins. Quant aux fusées, elles peuvent fort bien ne pas redescendre du tout. Bien sûr, la fusée peut éjecter un paquet d'instruments, qui redescend de son côté en parachute, mais ce n'est pas très sûr. Aussi, à elles seules, les fusées n'auraient pas permis grand-chose pour l'exploration de la haute atmosphère, si leur efficacité n'avait été décuplée par une autre invention : la *télémétrie*, appliquée pour la première fois à la recherche atmosphérique dans un ballon, en 1925, par le chercheur soviétique Pyotr A. Molchanoff.

Dans son principe, cette technique de « mesure à distance » consiste à traduire les mesures faites (la température par exemple) sous forme d'impulsions électriques qui peuvent être transmises à terre par radio. On observe alors les changements d'intensité ou d'espacement des impulsions reçues. Par exemple, une variation de température change la résistance électrique d'un fil, et donc la nature de l'impulsion envoyée par radio. Une variation de pression peut aussi être traduite par le changement de résistance électrique d'un fil, le refroidissement de celui-ci dépendant de la pression de l'air environnant. Les radiations reçues sont traduites par un détecteur approprié, et ainsi de suite. De nos jours, la télémétrie est devenue si perfectionnée qu'il semble ne plus manquer aux fusées que la parole, et que leurs messages complexes demandent, pour être déchiffrés, l'usage d'ordinateurs rapides.

Ainsi, les fusées et la télémétrie montrent qu'au-dessus de la stratosphère la température remonte jusqu'à un maximum de l'ordre de –10° C à une hauteur de 50 kilomètres, pour redescendre ensuite jusqu'à un minimum de –90° C à 80 kilomètres d'altitude. Cette région où la température monte, puis descend, s'appelle la *mésosphère*, nom proposé en 1950 par le géophysicien britannique Sydney Chapman.

Au-delà de la mésosphère, l'air raréfié ne forme plus que quelques millièmes de 1 % de la masse totale de l'atmosphère. Cette dispersion des molécules de l'air élève progressivement la température à une valeur de 1 000° C vers 500 kilomètres d'altitude, et probablement davantage encore plus haut. On appelle donc cette région la *thermosphère*, ou « sphère chaude », un curieux écho de la sphère de feu d'Aristote. Bien sûr la température, dans ce contexte-ci, n'a pas la même signification que dans le langage de tous les jours. Elle mesure seulement la vitesse des particules.

Au-dessus de 500 kilomètres, on atteint l'*exosphère* (ainsi nommée par Lyman Spitzer en 1949) qui s'étend jusqu'à 1 500 ou 2 000 kilomètres et se fond progressivement dans l'espace interplanétaire.

L'augmentation de nos connaissances sur l'atmosphère nous permettra peut-être un jour de modifier le temps au lieu de nous contenter d'en parler. On a déjà fait un petit pas dans cette direction. Au début des années quarante, les chimistes américains Vincent Joseph Schaefer et Irwing Langmuir ont constaté que des températures très basses pouvaient produire des noyaux de condensation, autour desquels se formaient des gouttes de pluie. En 1946, un avion a lâché de la neige carbonique dans une couche de nuages, de manière à y former d'abord des noyaux de condensation, puis des gouttes de pluie. Une demi-heure plus tard, il pleuvait. Plus tard, Bernard Vonnegut a obtenu de meilleurs résultats encore avec de l'iodure d'argent envoyé depuis le sol. Ces nouveaux « faiseurs de pluie » scientifiques peuvent mettre fin aux sécheresses – ou essayer, car il faut de toute façon que les nuages soient présents pour que l'on puisse les ensemencer. En 1961, des astronomes soviétiques ont réussi en partie à nettoyer un coin de ciel pour y observer une éclipse.

Les autres tentatives pour agir sur le temps ont des buts variés : ensemencer les ouragans pour diminuer leur violence, dissiper le brouillard, faire crever un nuage de grêle avant qu'il atteigne les cultures, etc. Les résultats, jusqu'ici, sont au mieux encourageants. De plus, en aidant les uns, on nuit aux autres : un fermier peut désirer la pluie, qui gêne au contraire la foire voisine. Les tentatives pour modifier le temps sont une source possible de procès. Aussi, on ne peut pas prévoir ce que l'avenir nous réserve en la matière.

D'autre part, les fusées ne servent pas seulement à l'exploration (bien que ce soit le seul usage mentionné au chapitre 3). Elles peuvent servir – et servent – aux besoins quotidiens de l'humanité. D'ailleurs, certaines formes d'exploration peuvent avoir une utilité immédiate. Un satellite en orbite n'est pas obligé de s'intéresser uniquement à l'espace : il peut tourner ses instruments vers la Terre. C'est ainsi que les satellites permettent de voir d'un seul coup d'œil une grande partie de la Terre et d'étudier dans son ensemble la circulation générale de l'air.

Le 1er avril 1960, les États-Unis ont lancé le premier satellite météorologique, *Tiros I* (*T*elevision *I*nfra*R*ed *O*bservation Satellite). En novembre, c'était le tour de *Tiros II*, qui a envoyé en dix semaines plus de vingt mille photos de grandes portions de la surface terrestre et de sa couverture de nuages, y compris un cyclone en Nouvelle-Zélande et un nuage à trombes dans l'Oklahoma. *Tiros III*, lancé en juillet 1961, a photographié dix-huit tempêtes tropicales, et en septembre il a montré le cyclone Esther en train de se former dans les Caraïbes, deux jours avant que les moyens traditionnels permettent de le repérer. Le satellite plus sensible *Nimbus I*, lancé le 28 août 1964, a été capable de photographier

les nuages pendant la nuit. Finalement, des centaines de stations automatiques de transmission des observations sont en service dans beaucoup de pays et les prédictions météorologiques sans le secours des satellites sont désormais devenues impensables. Certes, les prédictions ne sont pas encore sûres, mais on est loin des « devinettes » grossières qui en tenaient lieu il y a seulement vingt-cinq ans.

Les progrès les plus spectaculaires et les plus utiles sont relatifs à la détection et à la surveillance des cyclones. Ces tempêtes redoutables font beaucoup plus de dégâts que dans le passé, car le littoral est beaucoup plus peuplé qu'à l'époque de la Seconde Guerre mondiale, et si on ne pouvait pas connaître avec précision la position et les mouvements des cyclones, il est hors de doute que leurs dégâts, matériels et humains, seraient beaucoup plus grands qu'ils le sont (pour ce qui est du prix et de l'utilité du programme spatial, cette détection des cyclones par satellite permet d'économiser beaucoup plus que le prix du programme).

Les satellites ont d'autres usages terrestres. Dès 1945, l'écrivain anglais de science-fiction Arthur C. Clarke avait émis l'idée que les satellites pourraient servir de relais radio, permettant de traverser continents et océans, et qu'il suffirait pour couvrir le monde entier de trois satellites judicieusement placés. Rêve fou à l'époque, c'était une réalité quinze ans plus tard : le 12 août 1960, les États-Unis ont lancé *Echo I*, mince ballon de plastique aluminié, qui se gonfle dans l'espace jusqu'à un diamètre de trente mètres, et peut servir de réflecteur passif des ondes radio. L'un des promoteurs de ce succès était John Robinson Pierce, des laboratoires Bell, qui avait lui-même publié, sous un pseudonyme, des nouvelles de science-fiction.

Le 10 juillet 1962, *Telstar I* a été lancé aux États-Unis. Il ne se contentait pas de réfléchir les ondes : il les recevait, les amplifiait et les réémettait. Grâce à lui, les émissions de télévision ont traversé pour la première fois l'Océan (ce qui n'a malheureusement pas amélioré la qualité des programmes). Le 26 juillet 1963, *Syncom II* s'est installé sur une orbite à 36 000 kilomètres de la surface terrestre. Sa période était tout juste de vingt-quatre heures, de sorte qu'il est toujours resté au-dessus du même point de la Terre, au-dessus de l'océan Atlantique. *Syncom III*, placé de la même manière au-dessus de l'océan Indien, a transmis aux États-Unis les Jeux olympiques du Japon en octobre 1964.

Un satellite de communications encore plus perfectionné, *Early Bird*, a décolé le 6 avril 1965. Il a assuré 240 lignes téléphoniques et un canal de télévision. La même année, l'Union Soviétique a commencé elle aussi à envoyer des satellites de communications. Dans les années soixante-dix, la télévision, la radio et le téléphone sont devenus des réseaux mondiaux, grâce aux satellites. Du point de vue technologique, la Terre est devenue un monde uni, et les forces politiques qui s'opposent à cette évidence sont de plus en plus archaïques, anachroniques et dangereuses.

Évidemment, les satellites permettent de dresser la carte de la surface terrestre et d'étudier ses nuages. Ce qui est moins évident, mais tout aussi vrai, c'est qu'ils peuvent étudier aussi la couverture de neige, les mouvements des glaciers et la structure géologique à grande échelle. A partir de ces études géologiques, on peut repérer des régions susceptibles d'être pétrolifères. On peut étudier les récoltes et les forêts, et repérer les zones où la végétation est malade. On peut détecter les incendies de forêts et déterminer les besoins en matière d'irrigation. On peut étudier l'Océan,

ses courants et les déplacements de ses poissons. Cette évaluation des ressources terrestres est la meilleure réplique à ceux qui critiquent le « gaspillage » spatial au nom des problèmes plus urgents qui se posent sur la Terre. C'est souvent à partir de l'espace que ces problèmes peuvent trouver une solution.

Finalement, il y a en orbite de nombreux *satellites espions* conçus pour détecter les mouvements de troupes, les dépôts militaires, et ainsi de suite. Et il ne manque pas de gens pour envisager de faire de l'espace un nouveau théâtre de la guerre, pour mettre au point des *satellites tueurs,* destinés à abattre les satellites ennemis, ou pour placer en orbite des armes perfectionnées, capables de frapper plus vite que celles qui sont à terre. C'est le côté néfaste de la conquête de l'espace, encore que cela n'ajoute pas grand-chose à la capacité d'une guerre thermonucléaire de détruire complètement la civilisation.

Les deux superpuissances, États-Unis et Union Soviétique, proclament qu'elles « défendent la paix » en dissuadant l'adversaire de faire la guerre, chacune sachant que le déclenchement d'une guerre amènerait aussi bien sa propre destruction que celle de l'ennemi. Cette théorie relève de la folie meurtrière, car on n'a jamais empêché la guerre en multipliant le nombre et la « qualité » des armements.

Les gaz qui forment l'air

LA BASSE ATMOSPHÈRE

Jusqu'aux temps modernes, on a considéré l'air comme une substance simple et homogène. C'est au début du XVIIe siècle que le chimiste flamand Jan Baptist van Helmont y soupçonna la présence de plusieurs gaz, de natures chimiques différentes. Il étudia la vapeur dégagée par les jus de fruits en fermentation (le *gaz carbonique*) et y reconnut une substance nouvelle. En fait, il fut le premier à utiliser le mot *gaz,* qu'il forma peut-être en 1620 à partir du mot *chaos,* par lequel les anciens Grecs désignaient la substance originelle qui avait été la matière première de l'Univers. En 1756, le chimiste écossais Joseph Black étudia à fond le gaz carbonique et montra définitivement qu'il s'agissait d'un gaz différent de l'air. Il en détecta même un peu dans l'air atmosphérique. Dix ans plus tard, Henry Cavendish étudia un gaz inflammable qui était absent de l'air, et que l'on appela bientôt *hydrogène.* Ainsi se trouvait clairement démontrée la multiplicité des gaz.

C'est le chimiste français Antoine Laurent de Lavoisier qui réalisa le premier que l'air était un mélange de gaz. Vers 1770, il chauffa du mercure dans un ballon fermé et constata que ce métal se combinait à une partie de l'air pour former une poudre rouge *(oxyde de mercure),* tandis que les quatre cinquièmes de l'air restaient gazeux, aussi longtemps que l'on poursuivait le chauffage. Dans ce gaz restant, une chandelle ne pouvait pas brûler, et une souris ne pouvait pas vivre.

Lavoisier en conclut que l'air était un mélange de deux gaz. Le premier, qui se combinait au mercure dans son expérience et représentait un cinquième de l'air, était le gaz qui entretient la vie et la combustion : il

l'appela *oxygène*. Au second, il donna le nom d'*azote* (« pas de vie », en grec). En fait, les deux gaz furent aussi découverts indépendamment à peu près à la même époque, l'azote en 1772 par le physicien écossais Daniel Rutherford, et l'oxygène en 1774 par le pasteur anglais Joseph Priestley.

Ceci suffit à démontrer que l'atmosphère terrestre est unique dans le système solaire. A part la Terre, on connaît sept mondes ayant une atmosphère dans le système solaire. Jupiter, Saturne, Uranus et Neptune (les deux premiers à coup sûr, les autres probablement) ont une atmosphère d'hydrogène, avec une certaine quantité d'hélium. Mars et Vénus ont une atmosphère de gaz carbonique, avec un peu d'azote. Titan a une atmosphère d'azote, avec un peu de méthane. La Terre seule a une atmosphère composée de deux gaz avec des concentrations du même ordre, et c'est aussi la seule atmosphère où l'oxygène est un composant important. L'oxygène est un gaz très actif, et des considérations élémentaires de chimie laisseraient supposer qu'il doit se combiner à d'autres éléments et disparaître de l'atmosphère sous sa forme libre. Nous reviendrons sur ce point plus loin dans le même chapitre. Pour l'instant, revenons à la composition chimique de l'air.

Vers le milieu du XIXe siècle, le chimiste français Henri Victor Regnault avait déjà analysé des échantillons d'air venant de toutes les parties du monde, et montré que la composition de l'air était partout la même : l'air contient 20,9 % d'oxygène, et on admet que tout le reste (à part des traces de gaz carbonique) est formé d'azote.

L'azote est un gaz relativement inerte, c'est-à-dire qu'il ne se combine pas volontiers à d'autres substances. On peut pourtant l'y contraindre : en le chauffant avec du magnésium, on obtient du *nitrure de magnésium*, sous forme solide. Quelques années après la découverte de Lavoisier, Henry Cavendish essaya d'épuiser l'azote, en le combinant avec l'oxygène en présence d'une étincelle électrique : il échoua. Quoi qu'il fît, il ne pouvait pas se débarrasser d'une dernière bulle de gaz, représentant moins de 1 % de la quantité d'air au départ. Cavendish pensa qu'il s'agissait d'un gaz inconnu, encore moins réactif que l'azote. Mais tous les chimistes n'étaient pas des Cavendish, et il allait s'écouler un siècle avant que l'on détermine la nature de ce résidu gazeux de l'air.

En 1882, le physicien anglais Robert John Strutt (Lord Rayleigh) compara la densité de l'azote extrait de l'air avec celle de l'azote obtenu dans certaines réactions chimiques, et il constata, à sa grande surprise, que le premier était nettement plus dense. Se pouvait-il que l'azote extrait de l'air ne soit pas pur, mais contienne une petite proportion d'un autre gaz, plus dense ? Le chimiste écossais Sir William Ramsay aida Lord Rayleigh à approfondir la question. A ce moment-là, ils disposaient d'un instrument tout nouveau : la spectroscopie. Ils chauffèrent le petit résidu de gaz laissé après l'épuisement de l'azote, et ils examinèrent son spectre. Ils trouvèrent un nouveau système de raies brillantes, système qui n'appartenait à aucun élément connu. A ce nouvel élément, très inerte, ils donnèrent le nom d'*argon* (à partir d'un mot grec signifiant « inerte »).

L'argon à lui seul forme la quasi-totalité du petit résidu laissé quand on enlève à l'air l'oxygène et l'azote. Mais à côté de lui, on trouve des traces d'autres constituants (avec des concentrations de quelques millioniemes). Entre 1890 et 1900, Ramsay allait y identifier quatre autres gaz inertes : le *néon* (du grec « nouveau »), le *krypton* (« caché »), le *xénon* (« étranger ») et l'*hélium*, qui doit son nom au fait qu'on l'avait d'abord,

trente ans plus tôt, identifié dans le Soleil. Ces dernières années, le spectroscope infrarouge a révélé la présence de trois autres gaz : l'*oxyde nitreux* (le « gaz hilarant »), dont on ignore l'origine, l'*oxyde de carbone* (produit par la combustion incomplète de bois, charbon, essence, etc.) et le *méthane* ou « gaz des marais », libéré par la putréfaction des matières organiques. On estime que chaque année 45 millions de tonnes de méthane sont libérées dans l'atmosphère par les vents intestinaux du bétail et des autres gros animaux.

LA STRATOSPHÈRE

Jusqu'ici, nous avons parlé de la composition des couches les plus basses de l'atmosphère. Et la stratosphère ? Teisserenc de Bort pensait qu'elle pouvait contenir des quantités appréciables d'hydrogène et d'hélium, flottant sur les gaz plus denses des couches inférieures. Il se trompait. Dans les années trente, des aéronautes russes ont rapporté en ballon des échantillons d'air de la haute atmosphère, et on a constaté que cet air était composé, lui aussi, de quatre parties d'azote et d'une partie d'oxygène, comme celui des couches basses.

Mais il y avait des raisons de croire à l'existence de gaz inhabituels encore plus haut dans l'atmosphère, et l'une des plus fortes était l'existence de la lumière du « fond du ciel ». Il s'agit de la très faible illumination générale du ciel nocturne, même en l'absence de Lune. Cette lumière est si diffuse qu'on ne la décèle qu'avec les instruments les plus sensibles d'astronomie.

L'origine de cette lumière est restée mystérieuse pendant de longues années. A la fin des années trente, l'astronome Vesto Melvin Slipher y a détecté des raies spectrales déjà trouvées en 1964 par William Huggins dans les nébuleuses, et attribuées à un élément inconnu baptisé *nébulium.* Mais par des expériences de laboratoire, l'astronome américain Ira Sprague Bowen a montré que ces raies étaient émises par l'*oxygène atomique,* c'est-à-dire un gaz formé d'atomes isolés d'oxygène, au lieu de molécules (constituées chacune de deux atomes) comme dans l'oxygène normal. De la même façon, des raies spectrales bizarres émises par l'aurore boréale ont été finalement reconnues comme celles de l'azote atomique. Ces deux variétés anormales sont produites dans la haute atmosphère par des radiations solaires très énergétiques, capables de « casser les molécules » en atomes séparés – idée avancée en 1931 par Sidney Chapman. Heureusement, cette réaction absorbe ces radiations énergétiques, ou tout au moins affaiblit leur intensité, avant qu'elles atteignent la basse atmosphère.

Selon Chapman, la lumière du fond du ciel provient de la recombinaison, la nuit, des molécules cassées le jour par le soleil. En se recombinant, les deux atomes rendent une partie de l'énergie consommée par leur séparation, de sorte que cette lumière est en somme de la lumière solaire, retardée, très atténuée, et renvoyée d'une façon tout à fait particulière. En 1956 cette théorie a été confirmée directement par des expériences menées à la fois en laboratoire et dans la haute atmosphère (avec des fusées) sous la direction de Murray Zelikoff. Les spectroscopes emportés par les fusées ont enregistré les raies vertes de l'oxygène atomique avec une intensité particulièrement grande vers 100 kilomètres d'altitude. La

proportion d'azote atomique est plus faible, parce que la molécule d'azote est plus solide que celle de l'oxygène. Mais la lumière rouge de l'azote atomique est quand même appréciable, surtout vers 150 kilomètres d'altitude.

Dans la lumière du fond du ciel, Slipher avait aussi trouvé des raies étrangement proches de celles du sodium. Mais il semblait si invraisemblable que la haute atmosphère puisse contenir du sodium que l'on avait discrètement ignoré ces raies. Après tout, que ferait là le sodium, qui n'est pas un gaz, mais un métal très réactif, qui ne se présente jamais sur Terre hors de composés chimiques, surtout le *chlorure de sodium* (sel de cuisine) ? Mais en 1938, des physiciens français ont établi qu'il s'agissait bien des raies du sodium. Que ce fût vraisemblable ou non, il y avait donc du sodium dans la haute atmosphère. Ce sont encore les fusées qui ont permis les expériences décisives : leurs spectroscopes ont enregistré sans le moindre doute la lumière jaune du sodium, avec une intensité particulièrement forte à 80 kilomètres d'altitude. D'où vient ce sodium ? On ne le sait pas encore : peut-être des embruns de l'Océan, ou de la vaporisation des météores. Chose plus curieuse encore : on s'est aperçu en 1958 que le *lithium* – métal voisin du sodium et assez rare – contribuait lui aussi à la lumière du fond du ciel.

Entre autres expériences, l'équipe de Zelikoff a produit un « fond du ciel » artificiel : une fusée a libéré à haute altitude un nuage d'oxyde nitrique, ce qui a accéléré la recombinaison des atomes d'oxygène dans la haute atmosphère. Depuis le sol, on a pu observer facilement la luminosité qui en résultait. Une expérience analogue avec la vapeur de sodium a connu le même succès. Quand les Soviétiques ont lancé *Lunik III* vers la Lune, en octobre 1959, ils lui ont fait émettre un nuage de vapeur de sodium pour signaler de façon visible son arrivée sur orbite.

A des altitudes un peu plus faibles, il n'y a plus d'oxygène atomique, mais les radiations énergétiques du Soleil sont encore assez intenses pour provoquer la formation de la variété d'oxygène possédant trois atomes par molécule : l'*ozone*. C'est à 24 kilomètres que sa concentration est maximale. Même à cet endroit (baptisé *ozonosphère* en 1913 par le physicien français Charles Fabry), l'ozone n'a qu'une concentration de 0,25 millionième. Mais cette faible concentration suffit à absorber assez de lumière ultraviolette pour protéger la vie terrestre.

L'ozone se forme par combinaison d'un atome d'oxygène libre (oxygène atomique) avec une molécule d'oxygène normal (deux atomes). Mais c'est un corps instable. La molécule à trois atomes se casse facilement pour restituer la molécule à deux atomes, beaucoup plus stable, sous l'action de la lumière du Soleil, de l'oxyde nitreux qui existe naturellement, en quantités minimes, dans l'atmosphère, et aussi sous l'action d'autres corps.

C'est l'équilibre entre formation et destruction qui laisse en permanence dans l'ozonosphère la faible concentration dont nous avons parlé. C'est le bouclier qui protège la Terre des rayons ultraviolets du Soleil (qui détruiraient beaucoup des molécules délicates des tissus vivants), depuis que l'oxygène existe en grande quantité dans l'atmosphère.

L'ozonosphère n'est pas loin au-dessus de la tropopause, et varie en hauteur de la même manière, étant plus basse aux pôles qu'à l'équateur. D'autre part, l'ozone est plus abondant aux pôles qu'à l'équateur, où l'effet destructeur du Soleil est plus intense.

Ce serait dangereux si la technique humaine produisait quelque chose qui soit susceptible d'accélérer la destruction d'ozone dans la haute atmosphère, et d'affaiblir le bouclier anti-ultraviolet. Ceci entraînerait en effet une augmentation des radiations ultraviolettes au sol, ce qui rendrait plus fréquent le cancer de la peau, surtout chez les gens à peau claire. Selon certaines estimations, une diminution de 5 % du bouclier d'ozone donnerait 500 000 cas supplémentaires de cancer de la peau chaque année dans le monde entier. D'autre part, un accroissement des radiations ultraviolettes au sol pourrait aussi affecter le *plancton* (ensemble des êtres vivants microscopiques qui peuplent l'Océan), ce qui aurait des conséquences très graves, car le plancton est la base de la vie marine, et dans une certaine mesure de la vie terrestre.

Or, il est effectivement possible que la technique humaine affecte l'ozonosphère. Les avions sont de plus en plus nombreux à voler dans la stratosphère, et les fusées traversent l'atmosphère tout entière sur le chemin de l'espace. Ces véhicules déversent dans la haute atmosphère des produits de combustion qui pourraient fort bien accélérer la destruction de l'ozone. C'est un argument mis en avant au début des années soixante-dix contre la mise au point d'avions supersoniques commerciaux.

Les « bombes » à pulvérisateur peuvent aussi, de manière inattendue, jouer un rôle dangereux, comme on l'a découvert en 1974. Ces bombes utilisent du fréon (un gaz dont nous reparlerons au chapitre 9) sous pression pour pulvériser en fines gouttelettes le contenu de la bombe (brillantine, déodorants, purificateurs d'air, etc.). Par lui-même, le fréon est certainement ce qu'on peut imaginer de plus inoffensif comme gaz : incolore et inodore, inerte, il ne réagit sur rien, et ne saurait nuire directement à personne. On en répandait ainsi gaillardement 800 000 tonnes dans l'atmosphère chaque année, au moment où l'on s'est aperçu du danger que cela pouvait présenter.

Ce gaz, ne réagissant avec rien, se répand progressivement dans l'atmosphère, et finit par atteindre l'ozonosphère. Or, des expériences de laboratoire ont montré qu'il pourrait bien accélérer la destruction de l'ozone. En est-il de même dans les conditions de la haute atmosphère ? On n'en est pas sûr, mais le danger est trop grand pour être négligé. Depuis le début de la controverse à son propos, le fréon est beaucoup moins utilisé dans les bombes à pulvérisateur. Mais on l'utilise aussi, et en quantités beaucoup plus importantes, dans les réfrigérateurs et les appareils à conditionnement d'air, et là il n'est pas facile de lui trouver un remplaçant. Or, une fois formé, le fréon se répandra tôt ou tard dans l'atmosphère, mettant peut-être en danger notre bouclier d'ozone.

L'IONOSPHÈRE

L'ozone n'est pas le seul constituant de l'atmosphère à être beaucoup plus fréquent en haute altitude qu'au voisinage de la surface. D'autres expériences utilisant les fusées ont montré que les spéculations de Teisserenc de Bort, concernant l'existence de couches d'hydrogène et d'hélium, n'étaient pas fausses : il s'était seulement trompé sur leur emplacement. De 300 à 1 000 kilomètres, là où l'atmosphère devient « presque du vide », il y a une couche d'hélium, que l'on appelle maintenant l'*héliosphère*. C'est le ralentissement par frottement du satellite

Echo I qui a permis en 1961 au physicien belge Marcel Nicolet de déduire l'existence de l'héliosphère, existence confirmée par les analyses directes réalisées sur place par le satellite *Explorer XVII,* lancé le 2 avril 1963.

Au-dessus de l'héliosphère on trouve une couche encore plus ténue formée d'hydrogène, la *protonosphère,* qui s'étend peut-être jusqu'à 70 000 kilomètres avant de se fondre dans le « vide» interplanétaire.

Température élevée et radiations énergétiques ne se contentent pas de séparer les atomes ou de les assembler de nouvelles manières : elles peuvent arracher des électrons aux atomes, c'est-à-dire les *ioniser.* Ce qui reste de l'atome s'appelle un *ion,* et présente par rapport à l'atome la caractéristique de porter une charge électrique. Le mot *ion,* employé pour la première fois vers 1830 par l'Anglais William Whewell, vient d'un mot grec qui signifie « voyageur ». En effet, quand un courant électrique traverse une solution contenant des ions, ceux qui sont chargés positivement se déplacent dans un sens, et ceux qui sont chargés négativement dans l'autre.

C'est un Suédois, jeune étudiant en chimie, Svante August Arrhenius, qui a été le premier à suggérer, en 1884, que les ions étaient des atomes chargés, parce que c'était la seule explication du comportement des solutions conductrices. Cette idée, présentée dans sa thèse de doctorat, était si révolutionnaire que les examinateurs hésitèrent longtemps avant de la recevoir. A ce moment-là, on ignorait encore les particules chargées qui forment les atomes, et l'idée d'un atome chargé semblait ridicule. Arrhenius vit tout de même sa thèse acceptée, mais d'extrême justesse.

Quand on découvrit l'électron, juste avant 1900 (voir chapitre 6), la théorie d'Arrhenius devint brusquement plausible. Il reçut le prix Nobel de chimie en 1903 pour le travail qui lui avait pratiquement coûté sa thèse dix-neuf ans plus tôt (cette anecdote ressemble à un scénario invraisemblable, c'est vrai, mais l'histoire des sciences contient beaucoup d'épisodes à côté desquels les inventions hollywoodiennes les plus folles semblent banales).

La découverte des ions dans l'atmosphère ne se fit qu'après les expériences de Marconi sur la télégraphie sans fil. Quand il parvint, le 12 décembre 1901, à envoyer des signaux de Cornouailles à Terre-Neuve, à travers 3 800 kilomètres d'Océan, le monde scientifique fut stupéfait : les ondes radio se déplaçant en ligne droite, comment faisaient-elles pour contourner la courbure de la Terre et atteindre Terre-Neuve ?

Un physicien anglais, Oliver Heaviside, et un ingénieur électricien américain, Arthur Edwin Kennelly, ne tardèrent pas à suggérer que les signaux radio pouvaient se réfléchir vers la Terre sur une couche de particules électrisées située dans la haute atmosphère. La *couche de Kennelly-Heaviside,* comme on l'appelle maintenant, a été finalement localisée en 1922 par le physicien anglais Edward Victor Appleton. Celui-ci a constaté un curieux phénomène d'affaiblissement (« fading ») dans les transmissions radio. Il a expliqué cet affaiblissement par l'interférence entre deux signaux : celui qui vient directement de l'émetteur, et celui qui se réfléchit d'abord sur la couche électrisée de la haute atmosphère. Cette onde retardée n'étant pas en phase avec la première, l'interférence est partiellement destructrice, d'où l'affaiblissement.

Dès lors, ce n'était pas difficile de déterminer la hauteur de la couche réfléchissante. Il suffisait à Appleton d'envoyer des signaux de longueur d'onde telle que l'affaiblissement soit total, c'est-à-dire que les deux signaux arrivent en opposition de phase. A partir de la longueur d'onde utilisée et de la vitesse de propagation des ondes radio (déjà connue), il a pu calculer

la distance supplémentaire parcourue par l'onde qui s'était réfléchie. C'est ainsi qu'il a déterminé en 1924 l'altitude de la couche de Kennelly-Heaviside : un peu plus de 100 kilomètres.

Cet affaiblissement se produisait en général la nuit. En 1926 Appleton a constaté que juste avant l'aube, les ondes radio n'étaient pas réfléchies par la couche de Kennelly-Heaviside, mais par des couches plus hautes (que l'on appelle parfois maintenant *couches d'Appleton*), qui commencent à une altitude de 225 kilomètres (fig. 5.3).

Pour toutes ces découvertes, Appleton a reçu le prix Nobel de physique en 1947. Il a défini la région importante de l'atmosphère baptisée *ionosphère* en 1930 par le physicien écossais Robert Alexander Watson-Watt. Elle comprend ce que l'on a appelé plus tard la mésosphère et la thermosphère, et se divise en plusieurs couches. La région D s'étend de la stratopause jusqu'à la *couche D*, ou couche de Kennelly-Heaviside, située à une centaine de kilomètres de haut. Là commence la *région E*, zone intermédiaire relativement pauvre en ions, qui va jusqu'aux couches d'Appleton : la *couche F1* à 225 kilomètres et la *couche F2* à 320 kilomètres. La première est la plus riche en ions, la seconde ne devenant appréciable que pendant la journée. Au-dessus de ces couches s'étend la *région F*.

Ces couches réfléchissent et absorbent seulement les ondes longues utilisées pour les transmissions radio ordinaires. Les ondes plus courtes, comme celles qu'utilise la télévision, les traversent pratiquement en totalité. Par conséquent, la diffusion de la télévision est de portée limitée – à moins d'utiliser des satellites-relais, qui permettent aux émissions en direct de traverser océans et continents. Les ondes radio venues de l'espace (et produites par exemple par les étoiles émettrices) traversent aussi l'ionosphère, et heureusement : sinon, la radioastronomie n'existerait pas, au moins à partir de la surface terrestre.

C'est le soir que l'ionosphère est la plus efficace, à la suite de l'ensoleillement du jour. Au contraire, elle est la plus faible le matin, parce que la nuit ions et électrons se recombinent. Les orages solaires, qui augmentent le flux de particules et de radiations de haute énergie envoyé vers la Terre, renforcent et épaississent les couches ionisées. C'est alors que les aurores boréales apparaissent au-dessus de l'ionosphère. A ces moments-là, la transmission des ondes radio à grande distance sur la Terre devient difficile, voire impossible.

L'ionosphère est seulement l'une des régions entourant la Terre où se manifestent des phénomènes électriques. Hors de l'atmosphère, dans ce que l'on considérait comme le « vide », les satellites ont révélé dès 1958 l'existence de quelque chose de tout à fait étonnant. Pour comprendre de quoi il s'agit, il nous faut d'abord faire un détour, et parler du magnétisme.

Le magnétisme

Le nom vient de la ville grecque de Magnésie, près de laquelle ont été découverts, dès l'Antiquité, les premiers *aimants naturels*. Il s'agit d'un minerai (oxyde de fer) naturellement magnétique, dont la tradition veut qu'il ait été décrit pour la première fois par Thalès de Milet vers 550 av. J.-C.

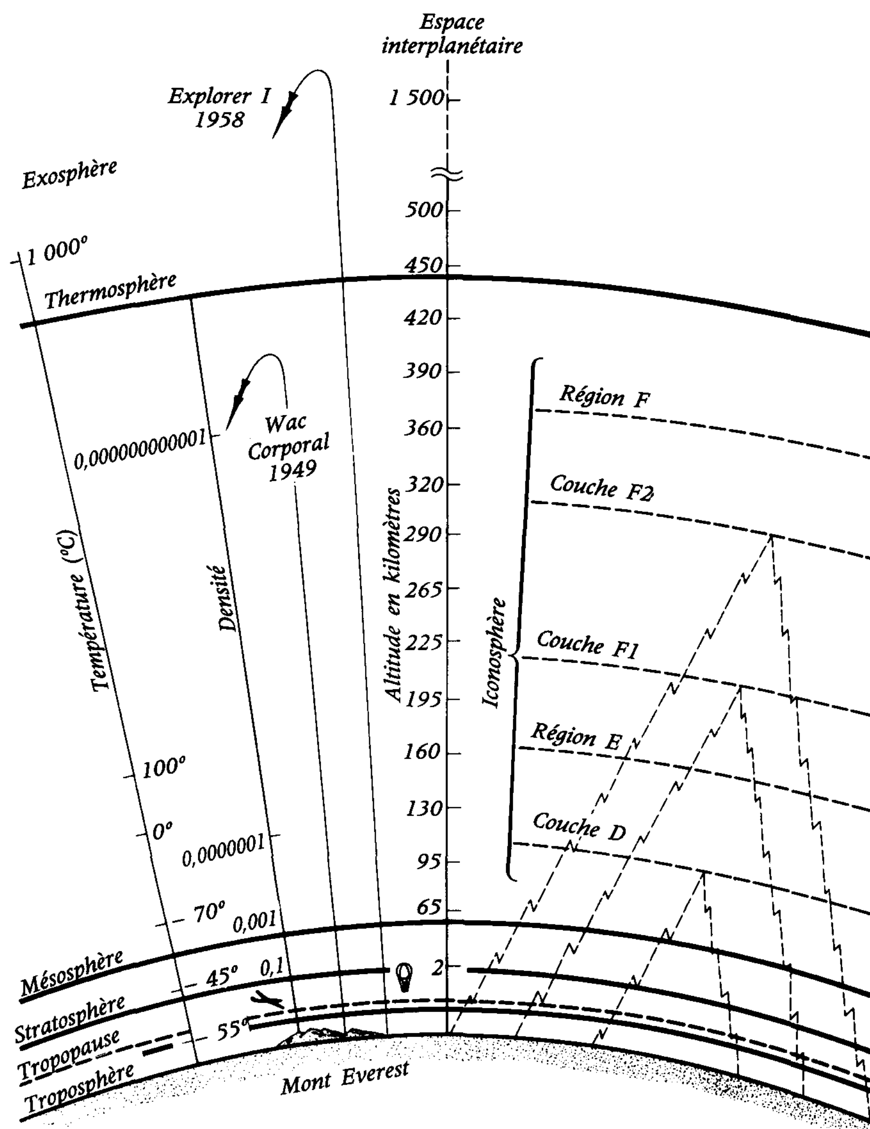


Figure 5.3. Coupe de l'atmosphère. Les lignes brisées indiquent les réflexions des ondes radio sur les couches de Kennelly-Heaviside et d'Appleton dans l'ionosphère. La densité de l'air décroît avec l'altitude (elle vaut 1 au niveau de la mer).

MAGNÉTISME ET ÉLECTRICITÉ

Les aimants acquirent une utilité pratique le jour où l'on découvrit qu'une aiguille de fer, frottée sur une « pierre d'aimant », devenait elle aussi un aimant, et que si on la montait sur un pivot vertical, elle s'orientait à peu près suivant une ligne nord-sud. Ceci était évidemment fort utile aux navigateurs et ne tarda pas à leur devenir indispensable – encore que les Polynésiens parvenaient à s'en passer pour de longs voyages d'île en île à travers le Pacifique.

On ne sait pas exactement qui réalisa la première boussole, mais il est généralement admis que c'est une invention chinoise, transmise aux Arabes, puis, par leur intermédiaire, aux Européens. En tout cas, la boussole se répandit en Europe au XIIe siècle, et en 1269 il en parut une description détaillée due à l'érudit français Petrus Peregrinus de Maricourt. C'est lui qui baptisa *pôle nord* le bout de l'aiguille qui pointait vers le nord, et *pôle sud* l'autre bout.

Bien sûr, on se demandait pourquoi une aiguille aimantée s'orientait de cette manière. Comme on savait que les aimants s'attiraient entre eux, certains pensaient qu'il existait dans le Nord le plus lointain une gigantesque montagne de « pierre d'aimant », vers laquelle se tournait l'aiguille de la boussole. (Cette montagne joue un grand rôle dans l'histoire de Sindbad le Marin, dans les *Mille et Une Nuits*.) D'autres, plus romantiques encore, attribuaient aux aimants une « âme » et une sorte de vie.

L'étude scientifique des aimants commença avec William Gilbert, le médecin de la reine Elisabeth I d'Angleterre. C'est lui qui découvrit que la Terre elle-même était un gigantesque aimant. Il monta une aiguille aimantée sur un pivot horizontal *(aiguille à inclinaison)* et son pôle nord se tourna vers le sol *(inclinaison magnétique)*. Gilbert tailla un aimant en forme de sphère et montra que l'aiguille se comportait de la même façon lorsqu'elle était placée au-dessus de l'hémisphère nord de cette « terre en miniature ». Il publia ses découvertes en 1600 dans un livre célèbre intitulé *De magnete*.

Pendant longtemps, les savants ont supposé que la Terre avait comme noyau un aimant gigantesque. Or, bien que l'on sache maintenant qu'elle a effectivement un noyau de fer, on sait aussi qu'il ne peut pas être aimanté, car le fer chauffé perd ses propriétés magnétiques *(ferromagnétisme)* à 760° C, et la température du noyau doit être supérieure à 1 000° C.

La température à laquelle une substance ferromagnétique perd cette propriété s'appelle *température de Curie*, car le phénomène a été découvert en 1895 par Pierre Curie. Le cobalt et le nickel, qui ressemblent étroitement au fer sous beaucoup d'aspects, sont également ferromagnétiques. La température de Curie du nickel est de 356° C, celle du cobalt de 1 075° C. D'autres métaux deviennent ferromagnétiques à basse température. Par exemple, le dysprosium est ferromagnétique au-dessous de — 188° C.

En général, le magnétisme est une propriété de l'atome lui-même. Mais dans la plupart des matériaux, les aimants minuscules que constituent les atomes sont orientés dans toutes les directions, et leurs effets s'annulent à peu près. Le faible magnétisme que l'on peut déceler s'appelle *paramagnétisme*. L'intensité du magnétisme s'exprime par la *perméabilité* du matériau : celle du vide est 1, et celle des substances paramagnétiques est comprise entre 1 et 1,01.

Les corps ferromagnétiques ont des perméabilités beaucoup plus élevées : 40 pour le nickel, 55 pour le cobalt, et plusieurs milliers pour le fer. Dans ces substances, selon une théorie présentée en 1907 par le physicien français Pierre Weiss, il existe des *domaines,* petits volumes dans chacun desquels les « aimants atomiques » sont alignés, ajoutant leurs effets pour produire des champs intenses. Ces domaines mesurent entre dix microns et un millimètre, et on les a effectivement observés depuis. Dans le fer non aimanté, les domaines sont orientés au hasard, et leurs effets s'annulent. Quand ils s'alignent sous l'effet d'un autre aimant, le fer est aimanté. Cette réorientation des domaines produit une sorte de crépitement, détectable par une amplification convenable. C'est ce que l'on appelle l'*effet Barkhausen,* du nom de celui qui l'a découvert, le physicien allemand Heinrich Barkhausen.

Dans les substances *antiferromagnétiques,* comme le manganèse, les domaines s'alignent aussi, mais en pointant tantôt dans un sens, tantôt dans l'autre, de sorte que l'effet global est presque nul. Au-dessus d'une certaine température, ces substances perdent leur antiferromagnétisme et deviennent paramagnétiques.

Si le noyau de fer de la Terre ne peut pas être lui-même un aimant permanent, parce qu'il est au-dessus de la température de Curie, il doit y avoir une autre explication à la déviation par la Terre de l'aiguille aimantée. Cette explication repose sur les travaux du physicien anglais Michael Faraday, qui a découvert la relation entre magnétisme et électricité.

Faraday a commencé, vers 1820, par une expérience très ancienne, déjà décrite par Petrus Peregrinus, et qui amuse toujours les élèves qui commencent à étudier la physique. Elle consiste à saupoudrer de limaille de fer un papier placé au-dessus d'un aimant, et à tapoter le papier. Les grains de limaille s'alignent suivant des arcs allant d'un pôle à l'autre de l'aimant. Faraday a décidé que ces arcs suivaient de véritables *lignes de force magnétiques,* qui forment un *champ magnétique.*

Or, en 1820, le physicien danois Hans Christian Oersted a remarqué qu'un courant électrique dans un fil faisait dévier une aiguille aimantée placée près de ce fil. Faraday en a conclu que le courant devait établir autour du fil un système de lignes de force magnétiques.

Il a été encouragé dans cette idée par les travaux du physicien français André Marie Ampère, qui s'est mis, dès la découverte d'Oersted, à étudier les fils parcourus par un courant. Ampère montra que deux fils parallèles, parcourus par le courant dans le même sens, s'attiraient mutuellement, et qu'ils se repoussaient quand les courants étaient de sens opposés. Ceci rappelle la manière dont deux pôles d'aimant de même nom se repoussent, tandis qu'un pôle nord et un pôle sud s'attirent. Mieux : Ampère montra qu'une bobine cylindrique de fil, parcourue par un courant, se comportait comme un barreau aimanté. C'est en l'honneur de ces travaux que l'on a donné, en 1881, le nom d'*ampère* à l'unité internationale d'intensité électrique.

Faraday, doté d'une intuition particulièrement pénétrante, s'est dit : si l'électricité crée un champ magnétique si semblable à celui des aimants que les bobines se comportent comme des aimants, pourquoi l'inverse ne serait-il pas possible ? Pourquoi ne pourrait-on pas, avec un aimant, créer un courant électrique identique à ceux que produisent les piles ?

En 1831, Faraday a réalisé une expérience qui allait changer l'histoire de l'humanité. Il a enroulé une bobine de fil autour d'un anneau de fer, et une autre bobine un peu plus loin sur le même anneau. Il s'est dit que, s'il faisait passer un courant dans la première bobine, ce courant allait créer des lignes de force magnétiques concentrées dans l'anneau, et que ce champ magnétique allait créer un courant dans la seconde bobine. Pour détecter ce courant, il a relié celle-ci à un *galvanomètre,* appareil de mesure du courant électrique inventé en 1820 par le physicien allemand Johann Salomo Christoph Schweigger.

Puis Faraday a relié la première bobine à une pile. Et ce qu'il attendait ne s'est pas produit : le courant avait beau passer dans la première bobine, rien ne passait dans la seconde. Mais Faraday a remarqué qu'à l'instant où il établissait le courant, l'aiguille du galvanomètre déviait brièvement, et à l'instant où il coupait le courant, l'aiguille déviait brièvement dans l'autre sens. Il réalisa aussitôt que c'était la variation du champ magnétique à travers la bobine – et non pas le champ magnétique lui-même – qui y faisait apparaître un courant. Quand il établissait le courant, le champ magnétique augmentait rapidement et engendrait un courant électrique qui durait autant que cette croissance. Inversement, quand il coupait le courant dans la première bobine, il apparaissait brièvement dans la deuxième un courant de sens opposé au premier.

Faraday a ainsi découvert le principe de l'induction électromagnétique, et accessoirement il a construit le premier *transformateur.* Il a imaginé aussitôt une manière plus simple et plus nette de mettre en évidence le même phénomène, en faisant entrer un aimant dans une bobine, et en l'en faisant sortir. En l'absence de toute pile, un courant traversait la bobine quand le champ magnétique à l'intérieur variait – ou, si on veut, quand le déplacement de l'aimant faisait « traverser » les fils par les lignes de force (figure 5.4).

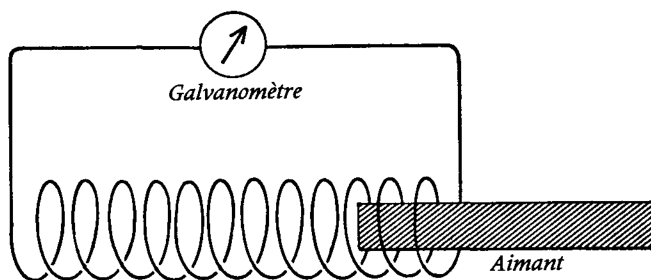


Figure 5.4. Une expérience de Faraday sur l'induction électromagnétique. Quand l'aimant se déplace dans la bobine, en entrant ou en sortant, un courant électrique apparaît dans le circuit.

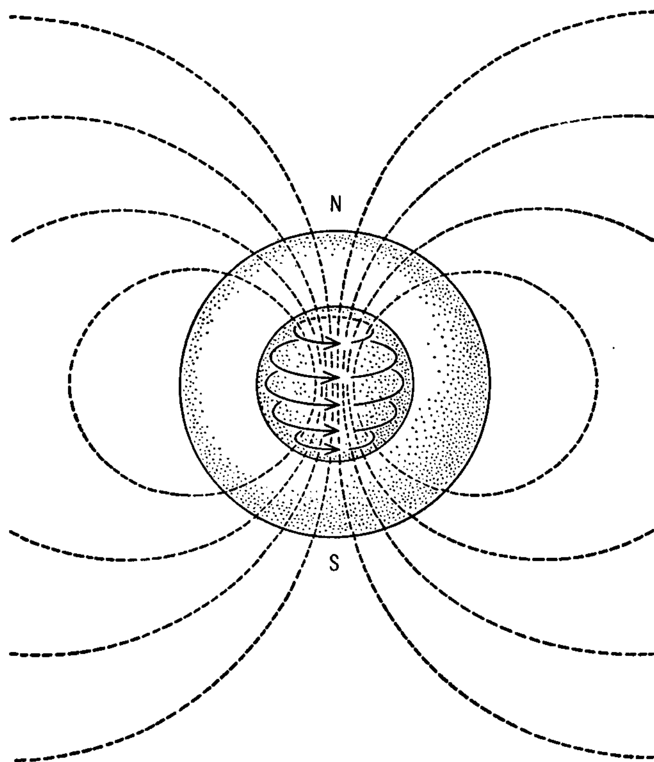
Non seulement les découvertes de Faraday allaient mener directement à la réalisation de la dynamo – et à la production de l'électricité –, mais elles sont à la base de la théorie électromagnétique de James Clerk Maxwell, qui a relié la lumière et les ondes radio (entre autres) pour former la famille des *ondes électromagnétiques.*

LE CHAMP MAGNÉTIQUE DE LA TERRE

Cette relation étroite entre le magnétisme et l'électricité suggère une explication possible du magnétisme terrestre, dont l'aiguille aimantée a permis de tracer les lignes de force : elles vont du *pôle magnétique nord* (au nord du Canada) au *pôle magnétique sud* (au bord de l'Antarctique), chacun se trouvant environ à 15° de latitude du pôle géographique correspondant. Bien sûr, l'aiguille aimantée ne donne que le champ au voisinage de la surface de la Terre, mais on a pu étudier ce champ à haute altitude grâce à des fusées emportant des *magnétomètres*. Quelle est donc l'explication proposée du magnétisme terrestre ? Il pourrait être créé par des courants électriques à l'intérieur de la Terre.

Selon le physicien Walter Maurice Elsasser, la rotation de la Terre provoque dans le noyau central de fer liquide des mouvements circulaires d'ouest en est. Ces mouvements produisent un courant électrique de même sens. Comme dans la bobine d'Ampère, ces courants circulaires créent un champ magnétique, identique à celui d'un gigantesque aimant disposé suivant l'axe des pôles. Cela donnerait au champ magnétique terrestre l'allure générale que l'on constate, avec des pôles voisins des pôles géographiques Nord et Sud (figure 5.5).

Figure 5.5. Origine du champ magnétique terrestre, selon la théorie d'Elsasser. Les déplacements de la matière dans le noyau liquide de fer et de nickel créent des courants électriques, qui engendrent le champ magnétique. Les lignes de force sont représentées en pointillés.



Le Soleil lui aussi possède un champ magnétique d'ensemble, deux ou trois fois plus fort que celui de la Terre, et des champs locaux, apparemment associés aux taches solaires, qui sont des milliers de fois plus intenses. En étudiant ces champs (grâce au fait qu'un champ magnétique intense modifie la longueur d'onde de la lumière émise), on se rend compte qu'il peut y avoir à l'intérieur du Soleil des mouvements circulaires de charges électriques.

En fait, les taches solaires présentent un certain nombre de bizarreries, que l'on ne comprendra qu'après avoir bien établi les causes des champs magnétiques à l'échelle astronomique. Au cours d'un cycle, les taches n'apparaissent qu'à certaines latitudes, qui varient pendant son déroulement. Le sens des champs magnétiques associés aux taches semble changer à chaque cycle, ce qui donne une période double (vingt et un ans) à ces variations de champ magnétique. On ignore encore les raisons de ces phénomènes.

Il n'y a d'ailleurs pas besoin d'aller jusqu'au Soleil pour rencontrer des énigmes magnétiques : la Terre en fournit déjà. Par exemple, pourquoi les pôles magnétiques ne coïncident-ils pas avec les pôles géographiques ? A chaque fois, la distance est environ de 1 500 kilomètres. De plus, les deux pôles magnétiques ne sont pas diamétralement opposés : la ligne qui les joint *(l'axe magnétique)* ne passe pas par le centre de la Terre.

En outre, l'angle entre l'aiguille aimantée et le vrai Nord (la direction du pôle Nord géographique) varie de façon irrégulière quand on voyage sur une ligne est-ouest. Christophe Colomb l'a constaté lors de son premier voyage et s'est bien gardé de le signaler à son équipage, déjà suffisamment terrifié et désireux de faire demi-tour.

C'est l'une des raisons pour lesquelles la boussole ne permet pas une détermination parfaite des directions. En 1911, l'inventeur américain Elmer Ambrose Sperry a proposé à cet effet une méthode non magnétique. Cette méthode est fondée sur la tendance d'une roue lourde en rotation rapide (un *gyroscope,* étudié d'abord par Foucault, celui qui avait démontré la rotation de la Terre) à résister aux modifications de son plan de rotation. On peut utiliser cette tendance pour réaliser un *gyrocompas,* qui permet de garder en référence une direction fixe et de guider navires ou fusées.

Mais si la boussole est imparfaite, elle a quand même été d'une utilité considérable pendant des siècles. On peut tenir compte de la déviation de l'aiguille par rapport au vrai Nord. Un siècle après Colomb, en 1581, l'Anglais Robert Norman réalisa la première carte indiquant la vraie direction indiquée par l'aiguille (la *déclinaison magnétique*) en différents points du globe. En réunissant les points où cette déclinaison a la même valeur, on obtient des *lignes isogoniques,* qui serpentent du pôle nord magnétique au pôle sud magnétique.

Malheureusement, il faut corriger périodiquement ces cartes, car en un même point la déclinaison magnétique varie avec le temps. Par exemple, à Londres, elle a varié de 32° en deux siècles ; en 1600 elle valait 8° est, et elle a lentement tourné vers l'ouest jusqu'à valoir en 1800 24° ouest. Depuis, elle varie en sens inverse, et en 1950 elle valait seulement 8° ouest.

L'inclinaison magnétique (l'angle de l'aiguille avec l'horizontale) varie aussi lentement au cours du temps en un point donné de la Terre. C'est pourquoi les cartes indiquant les lignes *isocliniques* (joignant les points de même inclinaison magnétique) doivent également être revues périodiquement. Qui plus est, l'intensité du champ magnétique terrestre (qui

augmente avec la latitude, et qui est trois fois plus grande près des pôles que dans les régions équatoriales) varie aussi avec le temps, ce qui oblige à réviser les cartes indiquant les lignes *isodynamiques* correspondantes.

Comme tout ce qui concerne le champ magnétique, son intensité globale change. Depuis quelque temps, elle diminue : depuis 1670, le champ magnétique terrestre a perdu 15 % de son intensité. Si cette diminution devait continuer au même rythme, le champ s'annulerait vers l'an 4000. Et ensuite ? Est-ce qu'il continuerait à diminuer, devenant négatif (c'est-à-dire avec le pôle nord dans l'Antarctique, et l'inverse) ? Autrement dit, est-ce que le champ magnétique terrestre oscille périodiquement ?

Une manière de voir si c'est possible consiste à étudier les roches volcaniques. Quand la lave se refroidit, les cristaux se forment dans l'alignement du champ magnétique terrestre. Dès 1906, le physicien français Bernard Brunhes avait remarqué que certaines roches étaient aimantées en sens inverse du champ magnétique terrestre actuel. Cette découverte est passée inaperçue à l'époque, parce qu'elle semblait dépourvue de signification. Mais maintenant il en va tout autrement ; ces roches nous montrent non seulement que le champ magnétique terrestre a changé de sens, mais qu'il l'a fait à de nombreuses reprises : neuf fois dans les quatre derniers millions d'années, à intervalles irréguliers.

La découverte la plus spectaculaire en ce domaine concerne le plancher océanique. S'il est vrai que la roche fondue suinte en permanence à travers la Grande Faille et s'étale, alors en s'écartant de cette faille vers l'est ou vers l'ouest, on doit trouver des roches solidifiées depuis des temps de plus en plus longs. Or, si on étudie leur alignement magnétique, on trouve bien des changements de sens par bandes entières parallèles à la faille, à des intervalles variant entre 50 000 et 20 millions d'années, et d'une façon symétrique par rapport à la faille. La seule explication rationnelle jusqu'ici implique *à la fois* l'élargissement du plancher et les renversements du champ magnétique terrestre.

Mais il est plus facile de s'assurer de la réalité de ces renversements que d'en trouver une explication.

En plus des variations à long terme, le champ magnétique varie d'un moment à l'autre de la même journée. Ceci suggère une influence du Soleil. De plus, il y a des *jours agités,* où l'aiguille de la boussole bouge plus vivement que d'habitude. On parle alors d'*orage magnétique.* Identiques aux orages électriques, ces phénomènes sont généralement accompagnés d'une augmentation des aurores boréales, comme l'a signalé dès 1759 le physicien anglais John Canton.

L'*aurore boréale* (nom proposé en 1621 par le philosophe français Pierre Gassendi) est un spectacle mouvant de traînées lumineuses qui se plient et se déplient, donnant une impression de splendeur irréelle. L'équivalent dans l'Antarctique s'appelle *aurore australe.* En 1741, l'astronome suédois Anders Celsius remarqua que l'aurore boréale avait quelque chose à voir avec le champ magnétique terrestre. Les traînées lumineuses semblent suivre les lignes de force du champ magnétique terrestre, et sont particulièrement visibles aux endroits où ces lignes se resserrent – près des pôles magnétiques. Pendant les orages magnétiques, on peut voir ces aurores aussi loin vers le sud que Boston ou New York.

Il est facile de comprendre ce qui cause les aurores. Une fois découverte l'ionosphère, on s'est rendu compte que quelque chose (probablement un rayonnement solaire d'un type ou d'un autre) donnait de l'énergie aux

atomes de la haute atmosphère et les transformait en ions chargés électriquement. La nuit, ces ions perdent leur charge et leur énergie, celle-ci se manifestant dans la lumière aurorale. C'est une variété spécialisée de « lumière du fond du ciel », qui suit les lignes de force magnétiques, et se concentre près des pôles magnétiques – ce à quoi l'on s'attend de la part d'ions chargés électriquement. (La lumière du fond du ciel ordinaire met en jeu des atomes non chargés, et par conséquent est indifférente au champ magnétique.)

LE VENT SOLAIRE

Qu'en est-il des jours agités et des orages magnétiques ? Le coupable est sans doute encore le Soleil.

L'activité des taches solaires semble engendrer les orages magnétiques. Comment un changement relativement petit à 150 millions de kilomètres peut-il affecter la Terre ? Ce n'est pas facile de l'imaginer, et pourtant cela doit être le cas, car les orages magnétiques sont particulièrement fréquents en période de grande activité des taches.

Le début de la solution s'est présentée en 1859, quand un astronome anglais, Richard Christopher Carrington, a vu un point lumineux apparaître sur la surface du Soleil, durer cinq minutes, puis s'effacer. C'était la première observation d'une *protubérance solaire.* Carrington l'a expliquée par la chute d'un gros météore sur le Soleil, et s'est dit que le phénomène devait être très rare.

Mais en 1889, George E. Hale a inventé le *spectrohéliographe,* qui permettait de photographier le Soleil en utilisant seulement la lumière correspondant à une bande spectrale donnée. Cette méthode détectait facilement les protubérances, montrait qu'elles étaient fréquentes, et qu'elles correspondaient aux régions riches en taches solaires. De toute évidence, ce sont des éruptions d'une énergie extraordinaire, mettant en jeu d'une manière ou d'une autre le phénomène qui produit les taches solaires (phénomène toujours inconnu). Quand la protubérance est près du centre du disque solaire, elle est tournée vers la Terre, et ce qu'elle éjecte se met en route vers la Terre. Les protubérances centrales sont toujours suivies d'orages magnétiques sur la Terre à quelques jours d'intervalle – le temps mis par des particules chassées du Soleil pour atteindre l'atmosphère terrestre. Dès 1896, le physicien norvégien Olaf Kristian Birkeland a signalé et décrit le phénomène.

Il ne manquait d'ailleurs pas de raisons d'affirmer que, quelle que fût l'origine des particules, la Terre était baignée dans un « nuage » qui s'étendait très loin dans l'espace. Les ondes radio engendrées par les éclairs voyagent à haute altitude le long des lignes de force du champ magnétique terrestre, ce qu'elles ne feraient pas en l'absence de particules électrisées. (Ces ondes, qui produisent dans les récepteurs des sifflements bizarres, avaient été découvertes par hasard durant la Première Guerre mondiale par le physicien allemand Heinrich Barkhausen.)

Il ne semblait pourtant pas que ces particules fussent émises par le Soleil en « bouffées » isolées. En 1931 Sydney Chapman a étudié la couronne solaire, et il a été frappé par son étendue. Lors des éclipses totales, nous ne voyons que la partie interne de la couronne. Pour Chapman, les particules chargées présentes au voisinage de la Terre font encore partie

de la couronne. En ce sens, la Terre évolue dans la frange très ténue de l'atmosphère solaire. Chapman a imaginé la couronne en expansion continuelle vers l'espace, avec un renouvellement adéquat à la surface du Soleil. Ce flot de particules émises constamment par le Soleil dans toutes les directions serait à l'origine des perturbations du champ magnétique terrestre.

Cette hypothèse a été largement confirmée dans les années cinquante par les travaux de l'astrophysicien allemand Ludwig Franz Biermann. Depuis le début du siècle on attribuait à la pression de radiation du Soleil le fait que la queue des comètes était toujours tournée à l'opposé du Soleil et s'accroissait au périhélie. Cette pression existe bien, mais Biermann a montré qu'elle était loin d'être suffisante pour produire la queue des comètes. Il fallait pour cela quelque chose de plus robuste, qui était presque à coup sûr un flux de particules chargées. Cette idée a été reprise par le physicien américain Eugene Norman Parker, qui s'est représenté un flux continu de particules, avec des bouffées plus fortes au moment des protubérances. En 1958, il a donné à ce phénomène le nom de *vent solaire.* On a eu très vite la confirmation directe de son existence, grâce aux satellites soviétiques *Lunik I* et *Lunik II,* qui sont partis vers la Lune en 1959 et 1960. En 1962, la sonde américaine *Mariner II,* en route vers Vénus, en a apporté une preuve supplémentaire.

Le vent solaire n'est pas un phénomène local. Il y a des raisons de penser qu'il reste assez dense pour être détectable au moins jusqu'à l'orbite de Saturne. Près de la Terre, la vitesse des particules du vent solaire va de 350 à 800 km/s, et elles mettent trois ou quatre jours pour aller du Soleil à la Terre. Le Soleil perd ainsi un million de tonnes par seconde, ce qui semble énorme à notre échelle, mais devient insignifiant à celle du Soleil : depuis sa naissance, il a perdu moins de 1 % de sa masse sous forme de vent solaire.

Il est bien possible que le vent solaire affecte notre vie de tous les jours. Outre son effet sur le champ magnétique, ses particules chargées peuvent, en traversant la haute atmosphère, influer sur les changements de temps au sol. S'il en est ainsi, l'étude des variations du vent solaire viendra s'ajouter à l'arsenal de moyens dont disposent les météorologues.

LA MAGNÉTOSPHÈRE

Les lancements de satellites ont révélé un effet imprévu du vent solaire. L'une des premières tâches assignées à ces satellites était de mesurer le rayonnement dans la haute atmosphère et l'espace proche, en particulier la densité des *rayons cosmiques* (particules chargées d'énergie particulièrement élevée). Quelle était l'intensité de ces rayons au-dessus du bouclier atmosphérique ? Pour le savoir, on a équipé les satellites de *compteurs Geiger* (inventés en 1907 par le physicien allemand Hans Geiger, et bien améliorés en 1908). Ces compteurs contiennent un gaz soumis à un voltage juste insuffisant pour y faire passer un courant. Quand une particule à haute énergie pénètre dans le gaz, elle ionise un de ses atomes et le projette sur d'autres atomes, créant de nouveaux ions, et ainsi de suite. Cette cascade d'ions peut transporter un courant : pendant une fraction de seconde, le courant passe à travers le gaz. Cette impulsion de courant est transmise par radio à la Terre, et on peut ainsi compter le nombre de particules

qui pénètrent dans le compteur en une seconde, ce qui permet d'estimer la densité des particules à l'endroit où il se trouve.

Quand les Américains ont réussi à mettre en orbite leur premier satellite, *Explorer I*, le 31 janvier 1958, son compteur a détecté à peu près les concentrations de particules prévues jusqu'à plusieurs centaines de kilomètres d'altitude. Mais plus haut (et *Explorer I* est monté jusqu'à 2 500 kilomètres) il y en avait moins que prévu. En fait, par moments, il n'y en avait pas du tout ! On aurait pu craindre une défaillance du compteur, mais *Explorer III*, lancé le 26 mars 1958 sur une orbite d'apogée 3 300 kilomètres, a donné les mêmes résultats, confirmés également par le satellite soviétique *Spoutnik III*, lancé le 15 mai 1958.

Le responsable du programme, James A. Van Allen de l'université d'Iowa et son équipe, ont proposé une explication du phénomène : les compteurs indiquaient zéro non pas parce qu'il y avait peu de particules, mais parce qu'il y en avait trop. Elles entraient trop vite et « aveuglaient » le compteur (comme nous sommes aveuglés par une lumière trop vive).

Le 26 juillet 1958, *Explorer IV* a emporté sur orbite des compteurs spécialement conçus pour mesurer des concentrations très importantes. L'un d'eux, par exemple, était entouré d'une mince couche de plomb, qui devait jouer le rôle de lunettes de soleil, et ne laisser pénétrer qu'une faible portion des particules. Cette fois, les compteurs ne se taisaient plus : l'idée de la saturation par des concentrations trop grandes était juste. *Explorer IV* est monté à 2 200 kilomètres et a mesuré des concentrations de particules beaucoup plus grandes que tout ce que l'on avait prévu.

En fait, les satellites Explorer n'avaient atteint que les couches les plus basses de cette région riche en particules. A l'automne 1958, les États-Unis ont lancé en direction de la Lune deux véhicules, *Pioneer I* et *Pioneer III*, qui se sont éloignés tous deux à une centaine de milliers de kilomètres de la Terre. Ils ont détecté deux bandes principales riches en particules, les *ceintures de Van Allen*, que l'on a appelées plus tard *magnétosphère*, pour harmoniser les noms donnés aux régions de l'espace entourant la Terre.

On a d'abord pensé que la magnétosphère était symétrique, comme une sorte de « gros beignet » entourant la Terre, et que les lignes de force magnétiques étaient elles aussi symétriquement disposées. Mais une fois encore, les satellites ont fourni des résultats inattendus. En 1963, en particulier, on a placé les satellites *Explorer XIV* et *Imp-I* sur des orbites très elliptiques destinées à les emmener si possible au-delà de la magnétosphère.

On a constaté alors (figure 5.6) que la magnétosphère avait une frontière nette, la *magnétopause*, repoussée vers la Terre par le vent solaire du côté tourné vers le Soleil, mais s'étendant à des distances considérables de l'autre côté. La magnétopause est à 65 000 kilomètres de la Terre du côte ensoleillé, mais sa « queue » de l'autre côté s'étend peut-être à un ou deux millions de kilomètres. En 1966 la sonde soviétique *Luna X*, qui a fait le tour de la Lune, y a détecté un faible champ magnétique qui est sans doute la « queue » de la magnétosphère terrestre.

Le piégeage de particules chargées le long de lignes de force magnétiques avait été prédit en 1957 par un scientifique amateur grec, né en Amérique, Nicholas Christofilos, qui gagnait sa vie en vendant des ascenseurs. Il avait envoyé ses calculs aux spécialistes de la question, mais personne n'y avait prêté attention (dans le domaine scientifique, comme ailleurs, les professionnels ont tendance à ignorer les amateurs). C'est seulement quand

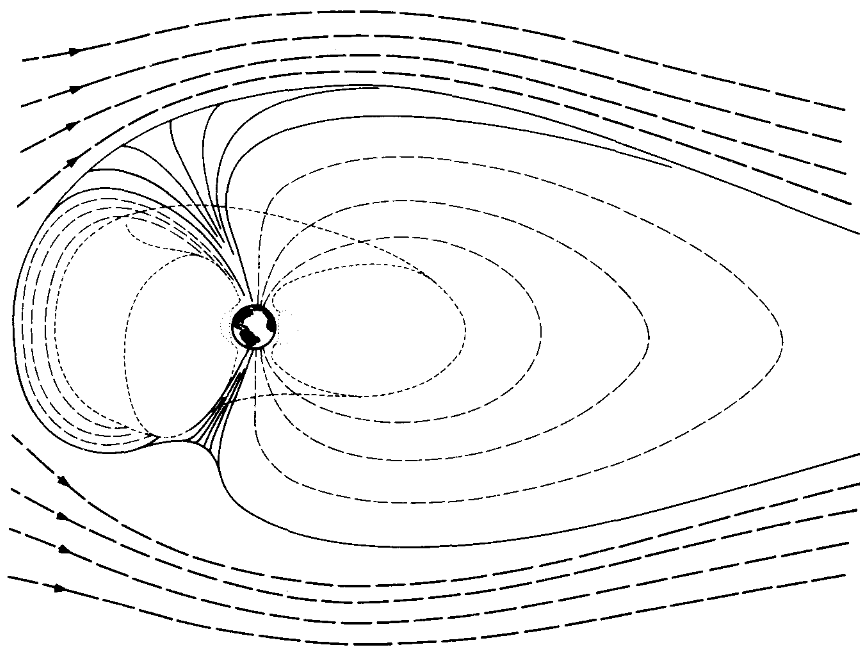


Figure 5.6. La magnétosphère (ou l'ensemble des ceintures de Van Allen) tracée par les satellites. Ces ceintures semblent formées de particules chargées, piégées par le champ magnétique terrestre.

les professionnels furent parvenus indépendamment aux mêmes résultats que l'on a reconnu les mérites de Christofilos, qui a été alors accueilli à l'université de Californie. Le piégeage magnétique des particules s'appelle maintenant *effet Christofilos.*

En août et septembre 1958, les États-Unis ont envoyé à 500 kilomètres de haut trois missiles nucléaires et les ont fait exploser, pour vérifier expérimentalement l'effet Christofilos – expérience baptisée programme Argus. Le flot de particules chargées produit par les explosiions s'est étalé le long des lignes de force et s'y est trouvé effectivement piégé. La bande qui en a résulté a persisté très longtemps : *Explorer IV* l'a détectée pendant plusieurs centaines de ses révolutions autour de la Terre. Ce nuage de particules a donné aussi des aurores – assez faibles – et perturbé pendant quelque temps les faisceaux radar.

Ce fut le début d'une série d'expériences affectant et même altérant l'espace au voisinage de la Terre, ce qui souleva l'indignation d'une bonne part de la communauté scientifique. Le 9 juillet 1962, une bombe atomique a explosé dans l'espace, provoquant dans la magnétosphère des changements apparemment durables, comme l'avaient prédit les scientifiques opposés à l'expérience, Fred Hoyle par exemple. Ces essais américains ont été suivis d'essais soviétiques similaires. Mais ces modifications artificielles peuvent nous empêcher de comprendre ce qui se passe naturellement dans la magnétosphère, et il est peu probable que ce genre d'expériences soit renouvelé.

Une autre perturbation – controversée elle aussi – a consisté à placer en orbite autour de la Terre une couche de fines aiguilles de cuivre pour mesurer leur capacité à réfléchir les ondes radio, et assurer ainsi une méthode sûre pour la communication à longue distance (l'ionosphère connaît de temps en temps des orages magnétiques pouvant interrompre les communications à un moment crucial). Les radioastronomes se sont élevés contre ce projet, susceptible de perturber les signaux reçus de l'espace. Mais ils n'ont pas été assez puissants pour empêcher sa réalisation : le programme West Ford (d'après la ville de Westford, Massachusetts, où ont eu lieu les préparatifs) s'est réalisé le 9 mai 1963. Un satellite a éjecté 400 millions d'aiguilles de cuivre, longues de deux centimètres et plus fines qu'un cheveu (le tout pèse 25 kilos). Ces aiguilles se sont étalées tout autour de la Terre en un anneau qui a persisté trois ans, et qui a bien réfléchi les ondes radio, mais pas assez pour être utilisables : il faudrait pour cela une couche beaucoup plus épaisse, et il est possible que, dans cette perspective, on tienne malgré tout compte de l'opposition des radio-astronomes.

LES MAGNÉTOSPHÈRES DES AUTRES PLANÈTES

Naturellement, on avait envie de savoir s'il y avait aussi des ceintures de particules autour des autres corps célestes. Si la théorie d'Elsasser est correcte, pour avoir une magnétosphère appréciable, un objet doit remplir deux conditions : un noyau conducteur liquide, où puissent s'établir des mouvements circulaires ; une période de rotation assez courte, pour pouvoir provoquer ces mouvements. La Lune, par exemple, a une densité trop faible pour posséder un noyau métallique, elle est trop petite et donc pas assez chaude au centre pour qu'un tel noyau soit liquide –, et de toute façon elle tourne sur elle-même trop lentement pour provoquer des mouvements dans un noyau liquide, s'il existait : il serait donc étonnant qu'elle possède une magnétosphère. Mais si convaincant que soit un raisonnement, il est toujours souhaitable d'en obtenir confirmation par des mesures directes, faciles à obtenir avec des sondes emportées par les fusées.

Effectivement, les sondes soviétiques *Lunik I* (janvier 1959) et *Lunik II* (septembre 1959) n'ont pas trouvé trace de ceintures de particules autour de la Lune, absence qui a été confirmée par toutes les sondes ultérieures.

Le cas de Vénus est plus intéressant. La planète est à peu près aussi massive et aussi dense que la Terre, et possède donc certainement un noyau métallique liquide. Mais elle tourne très lentement, encore plus lentement que la Lune. Toutes les sondes envoyées vers Vénus (depuis *Mariner 2* en 1962) ont confirmé que Vénus n'avait pratiquement pas de champ magnétique. Le peu qu'elle possède (moins d'1/20 000 de celui de la Terre) résulte peut-être de la pénétration, dans l'ionosphère, de son atmosphère très dense.

Mercure est dense aussi, et doit posséder un noyau métallique. Mais c'est encore une planète à rotation très lente. *Mariner 10,* qui a frôlé Mercure en 1973 et 1974, y a détecté pourtant un champ magnétique, faible, mais moins que celui de Vénus – et sans atmosphère pour l'expliquer. Si faible soit-il, ce champ est trop fort pour s'expliquer par la rotation très lente de la planète. Peut-être la petite taille de Mercure

(très inférieure à celle de la Terre ou de Vénus) permet-elle à son noyau d'être assez « frais » pour être ferromagnétique, et se comporter comme un aimant permanent ? Pour l'instant, on ne saurait le dire.

Mars tourne assez vite, mais c'est une planète plus petite et moins dense que la Terre. Toutefois, elle peut avoir un petit noyau métallique liquide, suffisant pour produire un effet appréciable. De fait, Mars semble posséder un petit champ magnétique, beaucoup plus faible que celui de la Terre, mais plus intense que celui de Vénus.

Avec Jupiter, tout change. Sa masse énorme et sa rotation très rapide en font un candidat de choix pour un champ magnétique élevé – à condition d'avoir un noyau conducteur. Or, en 1955, bien avant la construction des premières sondes spatiales, deux astronomes américains, Bernard Burke et Kenneth Franklin, ont détecté, en provenance de Jupiter, des ondes radio qui ne pouvaient pas résulter seulement d'effets de température. Elles avaient forcément une autre cause, peut-être des particules de haute énergie piégées dans un champ magnétique, comme l'a suggéré en 1959 Frank Donald Drake.

Avec les premières sondes jupitériennes, *Pioneer 10* et *Pioneer 11,* cette théorie allait être amplement confirmée. Il n'est pas difficile de détecter le champ magnétique de la planète géante : son intensité dépasse les estimations les plus optimistes. La magnétosphère de Jupiter est 1 200 fois plus étendue que celle de la Terre. Si elle était visible, elle remplirait dans notre ciel plusieurs fois la surface occupée par la pleine lune. La concentration des particules y est 19 000 fois plus grande que dans la magnétosphère terrestre. Ce serait un danger mortel pour des véhicules habités, un bouclier rendant la planète inapprochable – et englobant les satellites galiléens.

Saturne aussi a un champ magnétique intense – le deuxième, comme d'habitude... En l'absence d'observations directes, il semble raisonnable de supposer qu'Uranus et Neptune ont aussi des champs magnétiques plus intenses que le nôtre. Dans les planètes géantes gazeuses, quelle peut être la nature du noyau liquide conducteur ? Au moins dans le cas de Jupiter et de Saturne, il s'agit presque à coup sûr d'hydrogène métallique.

Météores et météorites

Les Grecs anciens savaient déjà que les « étoiles filantes » n'étaient pas vraiment des étoiles, parce que la population du ciel ne diminuait pas à la suite de leur chute. Pour Aristote, à une époque plus récente, le ciel était immuable et les étoiles filantes changeaient, donc elles ne faisaient pas partie du ciel, mais de l'atmosphère (comme quoi une idée fausse conduit parfois par hasard à des résultats corrects en partie). Aussi a-t-on donné à ces objets le nom de *météores* (qui veut dire « choses de l'air »). Ceux qui atteignent la surface de la Terre reçoivent le nom de *météorites.*

Les Anciens étaient parfois témoins de chutes de météorites ; ils constatèrent que certaines d'entre elles étaient des blocs de fer. On attribue à Hipparque un tel compte rendu. On suppose que la Kaaba, la pierre noire sacrée de La Mecque, est une météorite, qui doit sa sainteté à son

origine céleste. L'*Iliade* mentionne un bloc de fer brut parmi les prix des jeux à l'occasion des funérailles de Patrocle. Ce fer devait être météoritique, car à cette époque, l'âge du bronze, on ne connaissait pas encore la métallurgie du fer. Par contre, le fer météoritique était connu au moins dès l'an 3000 av. J.-C.

De la Renaissance au XVIIIe siècle, la science ne s'est pas intéressée aux météorites : elle s'attachait à combattre les superstitions et y rangea les « pierres venues du ciel ». Les paysans qui apportaient à l'Académie des sciences, à Paris, des échantillons de météorites se voyaient poliment mais fermement éconduits. Pourtant, en 1803, ces rapports incitèrent finalement le physicien Jean-Baptiste Biot à étudier sérieusement le phénomène. Ses recherches, scrupuleusement menées, ont contribué largement à convaincre le monde scientifique que des pierres tombaient effectivement du ciel.

Le monde scientifique, certes, mais pas le monde en général : en 1807, deux scientifiques du Connecticut (dont le jeune chimiste Benjamin Silliman) déclarèrent avoir assisté à la chute d'une météorite. Mais le président des États-Unis, Thomas Jefferson, préféra croire que deux professeurs mentaient, plutôt que d'admettre que des pierres tombaient du ciel.

Le 13 novembre 1833, les États-Unis reçurent une pluie de météores (du type appelé « léonides », car ils semblent provenir d'un point situé dans la constellation du Lion). Pendant plusieurs heures, le ciel fut le siège d'un véritable feu d'artifice, le plus brillant de tous les temps. Il ne semble pas que des météorites aient atteint le sol, mais le spectacle allait stimuler l'étude des météores, et les astronomes allaient s'y consacrer sérieusement pour la première fois.

Dès l'année suivante, le chimiste suédois Jöns Jakob Berzelius lança un programme d'analyse chimique des météorites. Ces analyses finirent par fournir aux astronomes des renseignements précieux sur l'âge du système solaire, et même sur la composition chimique d'ensemble de l'Univers.

LES MÉTÉORES

En notant les époques de l'année où leur densité augmentait, et les régions du ciel d'où ils semblaient parvenir, les observateurs de météores sont parvenus à établir les orbites de plusieurs nuages de météores. Ils se sont ainsi aperçus que les pluies de météores se produisent quand la Terre coupe l'orbite d'un nuage de météores.

Ceux-ci ont des orbites allongées, comme les comètes, ce qui incite à les considérer comme les débris de comètes désintégrées. La structure des comètes, selon la théorie de Whipple, les autorise à se désintégrer, laissant derrière elles de la poussière et des cailloux, et on a effectivement assisté à la désintégration de plusieurs comètes.

Quand un paquet de poussière de comète pénètre dans l'atmosphère, il peut faire un feu d'artifice aussi beau que celui de 1833 : un météore d'un gramme donne une « étoile filante » aussi brillante que Vénus. Certaines étoiles filantes visibles correspondent à des météores encore dix mille fois moins lourds !

On peut calculer le nombre total de météores qui atteignent l'atmosphère terrestre, et on s'aperçoit qu'il est extraordinairement élevé : chaque jour, il y en a plus de 20 000 qui pèsent au moins un gramme, et près de

200 millions d'autres plus légers, mais capables pourtant de donner une étoile filante visible à l'œil nu. Quant aux plus petits encore, ils se comptent chaque jour en milliards.

L'existence de ces *micrométéores* est révélée par la présence dans l'air de poussières de formes bizarres, contenant une forte proportion de nickel, et très différentes des poussières terrestres normales. Une autre preuve de leur existence est la *lumière zodiacale* (découverte vers 1700 par G.D. Cassini), ainsi nommée parce qu'elle est surtout visible au voisinage du plan de l'orbite terrestre, plan qui traverse les constellations du zodiaque. La lumière zodiacale est très faible, et ne peut être vue que par des nuits sans lune exceptionnellement transparentes. Elle est surtout nette près de l'horizon où le Soleil vient de se coucher, ou se prépare à se lever. Du côté opposé du ciel, on distingue une luminosité secondaire appelée Gegenschein (« lumière opposée », en allemand). La lumière zodiacale diffère de la lumière du fond du ciel : son spectre ne contient pas les raies de l'oxygène atomique ou du sodium, mais c'est seulement celui de la lumière solaire diffusée, rien de plus. Ce qui la diffuse, c'est sans doute de la poussière concentrée dans le plan des orbites planétaires – c'est-à-dire des micrométéores. A partir de l'intensité de la lumière zodiacale, on peut donc estimer leur nombre et leur taille.

On peut aussi les compter avec précision grâce à des satellites comme *Explorer XVI* (lancé en décembre 1962) et *Pegasus I* (lancé le 16 février 1965). Pour les détecter, certains satellites portent des feuilles d'une substance sensible qui signale chaque impact de météore par un changement de résistance électrique. D'autres enregistrent les impacts au moyen d'un microphone très sensible. Les résultats obtenus montrent que 3 000 tonnes de météores pénètrent chaque jour dans l'atmosphère, les cinq sixièmes étant formés de micrométéores trop petits pour se manifester sous forme d'étoiles filantes. Ces micrométéores forment un nuage ténu de poussières tout autour de la Terre, nuage qui s'étend, de moins en moins dense, jusqu'à 100 000 ou 200 000 kilomètres avant de se fondre dans la poussière normale de l'espace interplanétaire.

Grâce à la sonde vénusienne *Mariner 2*, on sait que la densité des poussières dans l'espace est dix mille fois plus faible qu'au voisinage de la Terre – qui apparaît ainsi comme le centre d'un nuage de poussières. Selon Fred Whipple, cela pourrait venir de la Lune, le nuage de poussières résultant du bombardement incessant de sa surface par les météorites. Vénus, qui n'a pas de lune, n'a pas non plus de nuage de poussières.

Le géophysicien Hans Petterson, qui s'est particulièrement intéressé à cette poussière météoritique, a recueilli des échantillons d'air en 1957 en haut d'une montagne de Hawaï, donc aussi loin des pollutions industrielles qu'il est possible sur la Terre. Il a pu ainsi estimer que la Terre reçoit chaque année cinq millions de tonnes de ces poussières (des mesures analogues, réalisées en 1964 par James M. Rosen avec des instruments emportés par ballons, aboutissent à une estimation de quatre millions de tonnes, mais certains auteurs se contentent de 100 000 tonnes par an). Hans Petterson a essayé d'estimer les variations passées de cette chute, en mesurant le contenu en nickel d'échantillons remontés du fond de l'Océan. Il semble y en avoir davantage dans les sédiments les plus superficiels, ce qui permet peut-être de supposer que la cadence du bombardement est en augmentation.

Les poussières météoritiques ont peut-être une grande influence sur notre vie : en 1953, le physicien australien Edward George Bowen a émis l'idée qu'elles étaient peut-être à l'origine des noyaux de condensation nécessaires à la formation des gouttes de pluie. Si cette théorie est correcte, la cadence du bombardement météoritique influe de façon décisive sur le régime des pluies sur la Terre.

LES MÉTÉORITES

De temps en temps, des morceaux de matière plus gros que des graviers, parfois même beaucoup plus gros, pénètrent dans l'atmosphère. Ils peuvent être assez gros pour survivre à l'échauffement provoqué par l'air (à des vitesses qui vont de 13 à 70 km/s) et pour atteindre le sol. Comme nous l'avons vu, on leur donne le nom de météorites. On pense généralement que ce sont des astéroïdes, de petits « earth grazers » qui ont frôlé la Terre de trop près...

La plupart des météorites trouvées sur le sol (on en connaît environ 1 700, dont 35 pèsent plus d'une tonne chacune) sont formées de fer, et on a longtemps cru que les météorites de fer étaient beaucoup plus nombreuses que les météorites rocheuses. C'est tout à fait faux, et dû simplement au fait qu'un morceau de fer se remarque beaucoup plus dans un champ pierreux qu'une pierre de plus – bien qu'une fois examinée, une pierre météoritique soit différente des autres pierres.

C'est en recensant les météorites que l'on avait vu tomber que les astronomes se sont aperçus qu'en réalité les rocheuses étaient près de dix fois plus nombreuses que les métalliques. On a en particulier trouvé beaucoup de météorites rocheuses dans le Kansas, ce qui paraît étrange, jusqu'à ce que l'on réalise que sur le sol sédimentaire, dépourvu de pierres, des champs du Kansas, une météorite rocheuse était aussi facile à repérer qu'un bloc de fer ailleurs.

Comment se forment ces deux types de météorites ? Les astéroïdes, dans la jeunesse du système solaire, ont pu être en moyenne plus gros que maintenant. Une fois formés, et rendus incapables de condensation supplémentaire par les perturbations dues à Jupiter, tout ce qui pouvait leur arriver était de se heurter entre eux, parfois assez fort pour voler en éclats. Mais auparavant, ils avaient pu devenir assez chauds, pendant leur formation, pour que leurs composants se séparent, le fer coulant vers le centre et la roche légère se rassemblant à l'extérieur. Aussi, une fois réduits en pièces, ces astéroïdes donnent deux sortes de morceaux, rocheux et métalliques, correspondant aux deux types de météorites que l'on trouve maintenant sur la Terre.

Il y a un troisième type de météorites, très rare, les *chondrites carbonées.* Une discussion à leur sujet sera plus à sa place au chapitre 13.

Les météorites font rarement des dégâts. Bien que 500 météorites de bonne taille tombent chaque année sur la Terre (dont malheureusement on retrouve seulement une vingtaine), la surface terrestre est vaste, et seules des régions limitées ont une densité de population importante. Pour autant qu'on sache, personne n'a jamais été tué par une météorite, bien qu'une habitante de l'Alabama ait raconté avoir été légèrement blessée par une météorite qui l'aurait frôlée, le 30 novembre 1955. En 1982, une météorite a traversé une maison à Westfield, Connecticut, sans atteindre

les habitants. Assez curieusement, Westfield avait déjà été atteint, sans mal, onze ans plus tôt.

Pourtant, les météorites peuvent avoir un effet dévastateur. En 1908, par exemple, un impact dans le centre de la Sibérie a creusé des cratères de cinquante mètres de diamètre et abattu les arbres à trente kilomètres à la ronde. Heureusement, il s'agissait d'une région déserte, et les seules victimes ont été des rennes. Si le choc s'était produit cinq heures plus tard, la rotation de la Terre aurait amené à cet endroit la ville de Leningrad (qui s'appelait alors Saint-Pétersbourg, et était la capitale de la Russie). Dans ce cas, la ville aurait été anéantie aussi totalement que par une bombe H. On estime que cette météorite pesait environ 40 000 tonnes.

Cette affaire de la *Toungouska* (ainsi nommée du nom de l'endroit) présente un certain nombre de mystères. Les difficultés d'accès de la région et la confusion engendrée peu après par la guerre et la révolution ont rendu impossible toute recherche sérieuse pendant des années. Quand on a pu étudier le site, il s'est révélé complètement dépourvu de matériaux météoritiques. Il y a quelques années, un auteur soviétique de science-fiction a construit une nouvelle sur l'idée que l'endroit était radioactif – idée immédiatement considérée comme un fait par tous ceux qui sont avides de sensationnel. Aussitôt on a avancé une foule d'hypothèses plus échevelées les unes que les autres, depuis l'impact d'un « mini-trou noir » jusqu'à l'accident d'un vaisseau extra-terrestre... L'hypothèse la plus plausible, malgré tout, est qu'il s'agissait d'un bloc de glace, probablement une toute petite comète, ou un morceau de comète (peut-être de la comète Encke), qui aurait explosé en l'air avant de toucher le sol, et aurait ainsi causé des dégâts considérables sans laisser aucune « signature » météoritique rocheuse ou métallique.

Depuis, le choc le plus important s'est produit en 1947, près de Vladivostok (encore en Sibérie).

Mais il y a des traces de chocs encore plus violents datant de l'époque préhistorique. Dans le comté de Coconino, en Arizona, il y a un cratère circulaire, d'environ 1 300 mètres de diamètre, et de 200 mètres de profondeur, entouré d'un bourrelet de terre de 30 à 50 mètres de haut. On dirait un cratère lunaire en miniature. On a longtemps supposé qu'il s'agissait d'un volcan éteint, mais un ingénieur des mines, Daniel Moreau Barringer, a soutenu qu'il s'agissait du résultat d'une chute de météorite, et le trou en question s'appelle maintenant cratère Barringer. Il est entouré de blocs de fer météoritique – des milliers, et peut-être des millions de tonnes en tout. Bien que l'on n'en ait extrait qu'une petite partie, c'est quand même plus que tout ce qu'on a trouvé de fer météoritique dans le reste du monde. Une preuve supplémentaire de l'origine météoritique du cratère est la présence, découverte en 1960, de formes de silex qui n'ont pu être produites que par les pressions énormes et les températures considérables entraînées momentanément par un tel impact.

Produit dans le désert il y a environ 25 000 ans par une météorite de fer de 50 mètres de diamètre, le cratère Barringer est très bien conservé. Dans la plupart des régions du monde, il aurait été estompé par le ruissellement et la végétation. Mais des observations aériennes révèlent parfois des formations circulaires ignorées jusque-là, en partie inondées et en partie recouvertes de végétation, et qui sont certainement d'origine météoritique. On en a découvert plusieurs au Canada, en particulier le cratère Brent dans le centre de l'Ontario, et le cratère Chubb dans le Nord

du Québec, chacun mesurant plus de trois kilomètres de diamètre, ainsi qu'au Ghana, où le cratère Ashanti dépasse 10 kilomètres de diamètre. Ces cratères ont peut-être plus d'un million d'années. On connaît dans le monde environ 70 de ces cratères fossiles, dont le diamètre dépasse parfois 130 kilomètres.

Les cratères de la Lune sont de tailles très diverses, depuis les petits trous jusqu'aux géants de 200 kilomètres de diamètre ou davantage. Privée d'air, d'eau et de vie, la Lune est un merveilleux « musée des cratères », où ils ne s'usent pas, si ce n'est par l'action extrêmement lente des changements de température qui résultent de l'alternance des jours et des nuits lunaires, longs à chaque fois de deux semaines terrestres. Peut-être la Terre serait-elle aussi grêlée que la Lune sans l'action cicatrisante du vent, de l'eau et de la croissance végétale.

On a d'abord attribué aux cratères lunaires une origine volcanique, mais leur structure n'est pas du tout celle des volcans éteints de la Terre. Dès 1890, l'idée de leur origine météoritique s'est imposée peu à peu.

Selon ce point de vue, les « mers » de la Lune, vastes étendues planes et circulaires relativement pauvres en cratères, seraient le résultat des impacts de météorites particulièrement grosses. Cette idée a trouvé une confirmation très nette en 1968 quand des satellites, placés en orbite autour de la Lune, ont manifesté des irrégularités de vol inattendues. Ces irrégularités ne s'expliquent que par l'existence de zones particulièrement denses de la surface lunaire, produisant une légère augmentation locale du champ gravitationnel, suffisante pour modifier de façon perceptible la trajectoire d'un satellite. Ces zones plus denses que la moyenne semblent correspondre aux mers lunaires et ont reçu le nom de *mascons* (pour « concentrations de masse »). On pense immédiatement que les grosses météorites de fer responsables des « mers » y sont enterrées et sont beaucoup plus denses que les roches de la croûte lunaire. Moins d'un an après leur découverte, on connaissait déjà une bonne douzaine de ces mascons.

Il ne faut pourtant pas se représenter la Lune comme un monde « mort » au point de vue du volcanisme. Le 3 novembre 1958, l'astronome soviétique N.A. Kozyrev a observé dans le cratère Alphonse un point rougeâtre (dès 1780, William Herschel avait déjà signalé des points rougeâtres sur la Lune). Les études spectroscopiques de Kozyrev ont montré une émission probable de gaz et de poussières. Depuis, on a observé d'autres rougeoiements sur la Lune, et il semble certain qu'une activité volcanique s'y manifeste épisodiquement. Lors de l'éclipse totale de Lune de décembre 1964, on a constaté que 300 cratères étaient plus chauds que leur voisinage – même s'ils n'étaient pas assez chauds pour émettre de la lumière.

Les mondes sans air, en général, qu'il s'agisse de Mercure ou des satellites de Mars, de Jupiter ou de Saturne, sont criblés de cratères rappelant le bombardement subi, il y a quelque quatre milliards d'années, au moment où les planéticules s'aggloméraient en planètes. Rien ne s'y est passé depuis qui ait pu estomper ces marques.

Vénus est pauvre en cratères, peut-être à cause de l'effet corrosif de son épaisse atmosphère. L'un des hémisphères de Mars aussi, sans doute parce que le volcanisme a recouvert la surface d'une couche relativement récente. C'est ce qui se passe sur Io, couvert de la lave de ses grands volcans. Quant à Europe, s'il n'a pas de cratères, c'est sans doute parce que les météorites y traversent la couche de glace et disparaissent dans le liquide

sous-jacent, tandis que la surface gèle à nouveau, cicatrisant rapidement le trou.

Étant les seuls échantillons de matière extra-terrestre que nous puissions examiner, les météorites passionnent non seulement les astronomes, les géologues, les chimistes et les métallurgistes, mais aussi les cosmologistes, qui se penchent sur les origines de l'Univers et du système solaire. Or, parmi les météorites, on trouve en plusieurs points du globe de curieux objets vitreux. On a découvert les premiers en 1787 dans ce qui est maintenant l'Ouest de la Tchécoslovaquie, et les suivants en Australie en 1864. On les appelle *tectites* (à partir d'un mot grec signifiant « fondu »), car ils semblent avoir fondu en traversant l'atmosphère.

En 1936, l'astronome américain Harvey Harlow Ninninger a émis l'idée que les tectites pouvaient être des restes de la matière expulsée de la Lune par l'impact de grosses météorites et capturée par le champ gravitationnel de la Terre. On en trouve beaucoup en Australie, en Asie du Sud-Est et dans l'océan Indien. Ce groupe de tectites est le plus récent de tous : il n'a que 700 000 ans. Il est concevable qu'il ait pu être produit par l'impact qui a créé sur la Lune le plus jeune des grands cratères, Tycho. Le fait que cet impact ait coïncidé avec le dernier renversement du champ magnétique terrestre permet peut-être d'expliquer, par d'autres catastrophes mettant en jeu la Terre et la Lune, l'irrégularité frappante de ces changements de sens.

Autre groupe étonnant de météorites : celles de l'Antarctique. D'abord, on remarque immédiatement, sur la calotte glaciaire, n'importe quelle météorite, qu'elle soit rocheuse ou métallique. En fait, n'importe quel objet solide qui ne soit pas de la glace et qui ne soit pas d'origine humaine, sur ce vaste continent, est forcément d'origine météoritique. Et une fois sur le sol, plus rien ne lui arrive (au moins depuis vingt millions d'années) à moins qu'il ne soit enseveli sous la neige ou qu'un manchot empereur ne vienne y trébucher.

Il n'y a jamais beaucoup de présence humaine dans l'Antarctique, et on n'a étudié en détail qu'une faible partie de sa surface. Aussi il n'est pas étonnant que, jusqu'en 1969, on n'y ait trouvé que quatre météorites, et à chaque fois par hasard. Mais en 1969 une équipe de géologues japonais a trouvé neuf météorites très rapprochées. Cela a soulevé l'intérêt scientifique général, et on en a trouvé de plus en plus. En 1983, on a recensé sur le continent gelé plus de 5 000 fragments météoritiques – nettement plus que dans tout le reste du monde (encore une fois, cela ne veut pas dire qu'il en tombe davantage à cet endroit, mais seulement qu'ils y sont beaucoup plus faciles à repérer).

Or, certains de ces échantillons sont vraiment bizarres. En janvier 1982, on a découvert un fragment météoritique verdâtre qui, à l'analyse, s'est révélé remarquablement semblable aux échantillons de pierres lunaires rapportés par les cosmonautes. On imagine mal comment une pierre lunaire pourrait être projetée jusqu'à la Terre, et pourtant cela semble bien être le cas.

Ensuite, certains fragments météoritiques de l'Antarctique libèrent, quand on les chauffe, des gaz dont la composition rappelle étroitement celle de l'atmosphère de Mars. Qui plus est, ils semblent âgés seulement de 1,3 milliard d'années, au lieu de 4,5 milliards comme les météorites ordinaires. Or, il y a 1,3 milliard d'années, les volcans de Mars ont pu être particulièrement actifs. Peut-être certains de ces fragments sont-ils des morceaux de lave martienne projetés jusqu'à la Terre.

Remarquons en passant que l'âge des météorites (calculé par des méthodes que nous décrirons au chapitre 7) est un outil important de détermination de l'âge de la Terre et du système solaire en général.

L'air : l'acquérir et le garder

Avant de nous demander comment la Terre a pu acquérir son atmosphère, nous devrions peut-être chercher à savoir comment elle a pu s'y cramponner pendant ses milliards d'années de voyage dans l'espace. La réponse à cette question fait intervenir quelque chose qui s'appelle la *vitesse de libération.*

LA VITESSE DE LIBÉRATION

Si on lance un objet vers le haut, à partir du sol, l'attraction gravitationnelle ralentit son mouvement, jusqu'à ce qu'il s'arrête, rebrousse chemin et retombe. Si la gravité était la même quelle que soit la hauteur, l'altitude atteinte par l'objet serait proportionnelle au carré de sa vitesse de départ : il monterait à une certaine altitude avec une vitesse de départ verticale de 1 km/s, et à une altitude quatre fois plus grande avec une vitesse de départ de 2 km/s.

Mais en réalité, l'intensité de la pesanteur ne reste pas constante : elle diminue lentement avec l'altitude (elle est inversement proportionnelle au carré de la distance au centre de la Terre). Si nous lançons un objet vers le haut, avec une vitesse de départ verticale égale à 1 km/s, il va monter à 50 kilomètres d'altitude (en négligeant la résistance de l'air). Mais si nous le lançons avec une vitesse double, il va monter plus de quatre fois plus haut : à 50 kilomètres d'altitude, la gravité est déjà nettement plus faible qu'au sol, et au-dessus de cette altitude l'objet sera moins freiné que si la gravité était constante. En fait, au lieu de monter à 200 kilomètres, il montera à près de 220 kilomètres.

Avec une vitesse de départ verticale de 10,5 km/s, un projectile s'élève à 45 000 kilomètres. A cet endroit, la force de gravitation est quarante fois moins forte qu'à la surface de la Terre. Il suffit d'augmenter la vitesse initiale d'un peu plus de 100 m/s pour que le projectile atteigne 55 000 kilomètres d'altitude.

On peut montrer qu'un projectile lancé avec une vitesse de départ verticale de 11,2 km/s ne retombe plus : bien que la gravitation terrestre le freine toujours, ce freinage diminue avec la distance, et ne parvient plus à annuler la vitesse du projectile par rapport à la Terre (faisant mentir l'adage suivant lequel « tout ce qui monte finit par redescendre »).

Cette vitesse de 11,2 km/s est la *vitesse de libération* de la Terre. On peut calculer la vitesse de libération de n'importe quel corps céleste, connaissant sa masse et son rayon. Pour la Lune, elle vaut seulement 2,4 km/s, pour Mars 5 km/s, pour Saturne 36 km/s, et pour Jupiter, le géant du système, 60 km/s.

Or, cela détermine directement la capacité, pour la Terre, de retenir autour d'elle son atmosphère. En effet, les molécules et les atomes de l'air

filent constamment dans tous les sens comme de minuscules projectiles. Leur vitesse varie considérablement, mais on peut en calculer certaines valeurs statistiques, par exemple la vitesse moyenne dans certaines conditions, ou la proportion des molécules ayant à un instant donné une vitesse supérieure à une certaine valeur. La méthode qui permet de le faire a été mise au point en 1860 par James Clerk Maxwell et le physicien autrichien Ludwig Boltzmann, et se traduit par la *loi de Maxwell-Boltzmann*.

On trouve ainsi que la vitesse moyenne des molécules d'oxygène, dans l'air à la température ambiante, vaut environ 500 m/s. La molécule d'hydrogène, seize fois moins lourde, va en moyenne quatre fois plus vite, soit 2 km/s (à une température donnée, le produit de la masse par le carré de la vitesse a la même valeur moyenne pour toutes les molécules).

Mais ce sont là des vitesses moyennes : la moitié des molécules va plus vite que cela, une certaine proportion va deux fois plus vite au moins, etc. En fait, une petite fraction des molécules d'oxygène et d'hydrogène de l'atmosphère va plus vite que 11,2 km/s, la vitesse de libération.

Dans la basse atmosphère, ces « bolides » ne peuvent pas s'échapper, parce que les collisions avec leurs voisines ne tardent pas à les ralentir. Mais dans la haute atmosphère, leurs chances d'évasion sont bien meilleures. Tout d'abord, là-haut, le rayonnement solaire est très intense et leur communique une grande énergie, et donc des vitesses énormes. Ensuite, la probabilité d'une collision est très réduite dans l'air ténu. Au voisinage du sol, une molécule voyage en moyenne un centième de micron avant d'en heurter une autre. Mais à 100 kilomètres d'altitude le « libre parcours moyen » entre deux collisions est de l'ordre de 10 centimètres, et deux fois plus haut il dépasse un kilomètre ! Là, un atome ou une molécule subit en moyenne un choc par seconde – au lieu de cinq milliards par seconde comme au niveau de la mer. Ainsi, une particule rapide à 200 kilomètres d'altitude a de bonnes chances de s'échapper de l'atmosphère. Si par hasard sa vitesse est dirigée vers le haut, elle pénètre dans des régions de plus en plus ténues, où la probabilité d'une collision diminue encore : elle peut finir par se retrouver dans l'espace interplanétaire, définitivement libérée de la Terre.

Autrement dit, l'atmosphère « fuit ». Mais cette fuite concerne surtout les molécules les plus légères. L'oxygène et l'azote ont des molécules assez lourdes pour qu'une petite fraction seulement puisse atteindre la vitesse de libération, et l'atmosphère a perdu très peu de ces gaz depuis leur formation. Au contraire, l'hydrogène et l'hélium atteignent facilement la vitesse de libération : il n'est pas surprenant qu'il n'en reste pratiquement pas dans l'atmosphère terrestre actuelle.

Les planètes plus massives, comme Jupiter et Saturne, peuvent retenir même l'hélium et l'hydrogène, ce qui leur permet d'avoir des atmosphères épaisses composées principalement de ces deux éléments (qui sont après tout les plus répandus dans l'Univers). L'hydrogène, en grande quantité, peut réagir avec les autres éléments qui sont présents, de sorte que le carbone, l'azote et l'oxygène n'existent que sous forme de composés hydrogénés, respectivement le méthane (CH_4) l'ammoniac (NH_3) et l'eau (H_2O). L'ammoniac et le méthane présents dans l'atmosphère de Jupiter, bien que leur concentration soit relativement faible, ont été découverts dès 1931 (par l'astronome germano-américain Rupert Wildt), parce qu'ils donnent des bandes d'absorption nettes dans le spectre, contrairement à l'hydrogène et à l'hélium. La présence de ces derniers n'a pu être révélée

II. La Terre et les voyages spatiaux

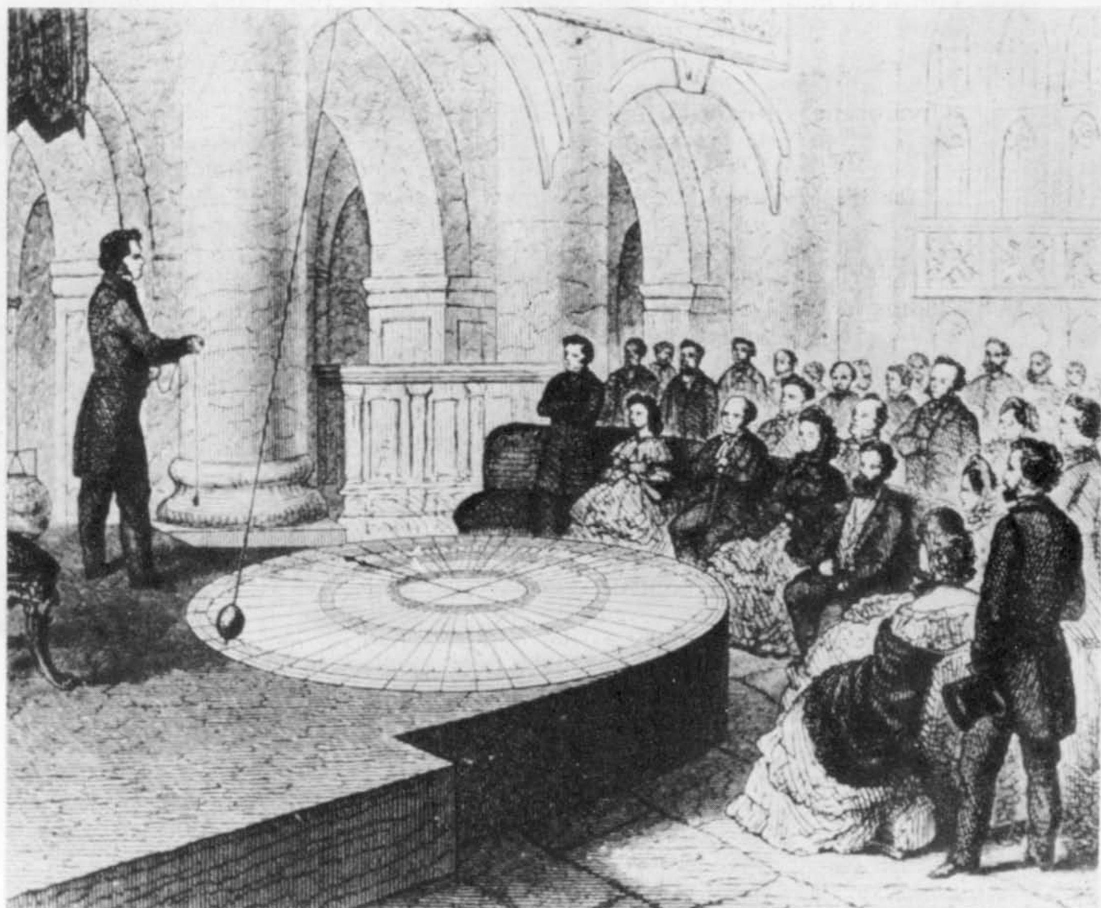


Planche II.1. La célèbre expérience de Foucault à Paris en 1851 démontre la rotation de la Terre autour de son axe au moyen d'un pendule, dont le plan d'oscillation tourne dans le sens des aiguilles d'une montre. (Avec l'aimable autorisation des archives Bettmann.)

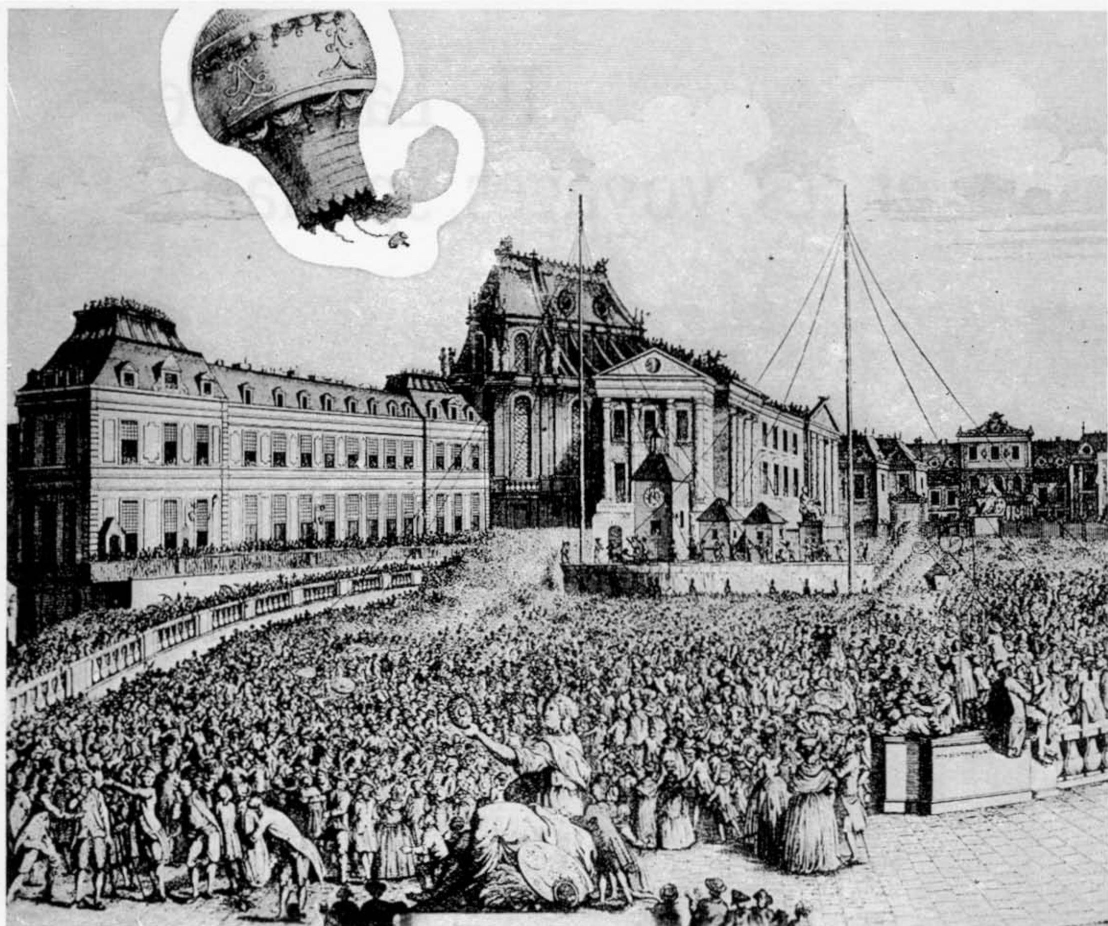


Planche II.2. Le ballon à air chaud des frères Montgolfier, lancé à Versailles le 19 septembre 1783. (Avec l'aimable autorisation des archives Bettmann.)



Planche II.3. Lancement du premier satellite américain, *Explorer I,* le 30 janvier 1958. (Avec l'aimable autorisation de l'armée américaine.)

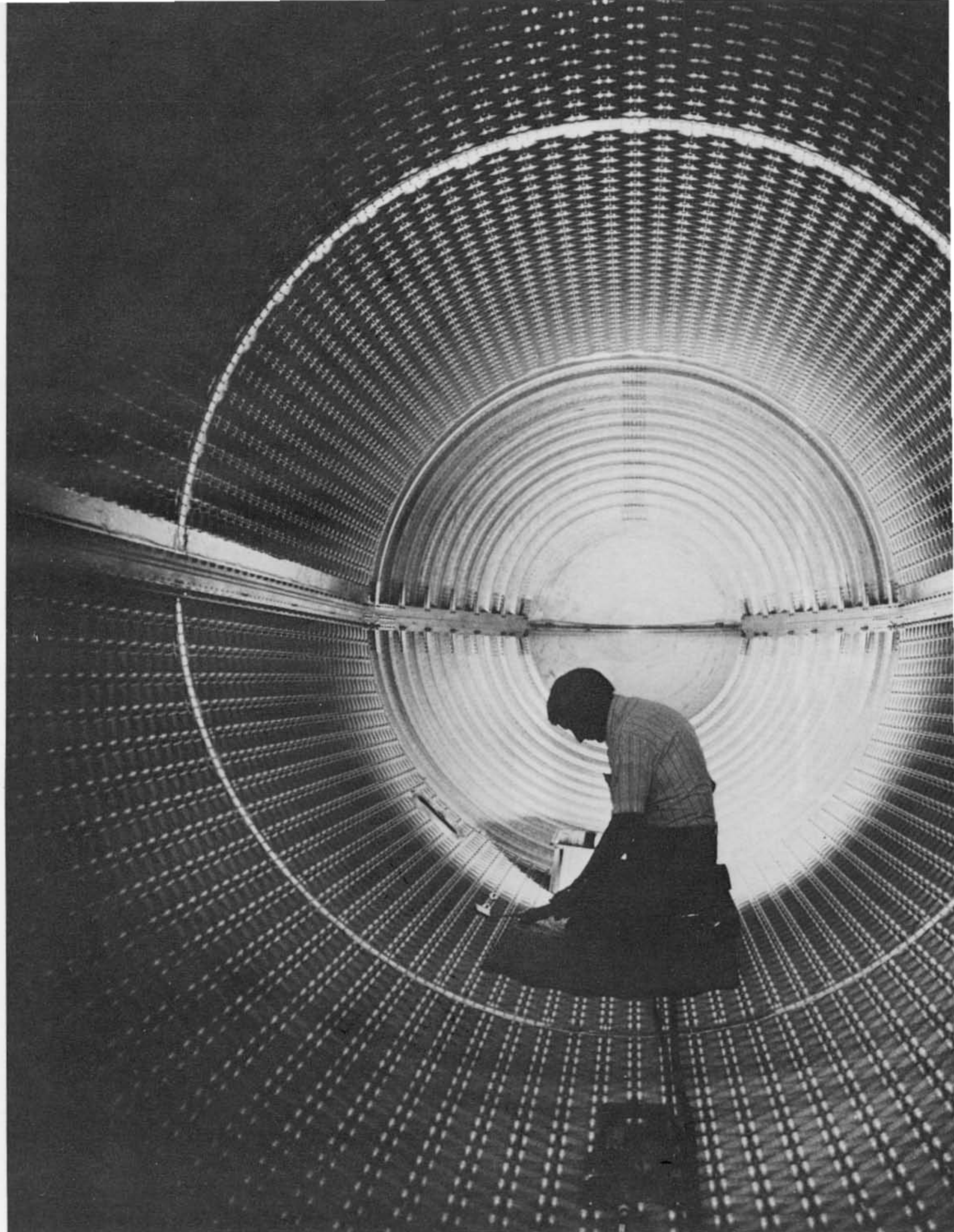


Planche II.4. Cette silhouette agenouillée devant un halo lumineux est celle d'un mécanicien au travail à l'intérieur du capot d'une fusée McDonnell Douglas. Sa lampe portative illumine les milliers de facettes du motif triangulaire qui permet de réduire le poids de ce capot d'aluminium brillant, tout en gardant une rigidité maximale. Ce capot, d'un diamètre de 2,50 mètres et d'une longueur de 8 mètres, protège la charge utile de la fusée *Delta* pendant son lancement et pendant la traversée de l'atmosphère, qui l'expose à des forces aérodynamiques et à un échauffement dangereux. (Avec l'aimable autorisation de la société McDonnell Douglas Astronautics, Californie.)

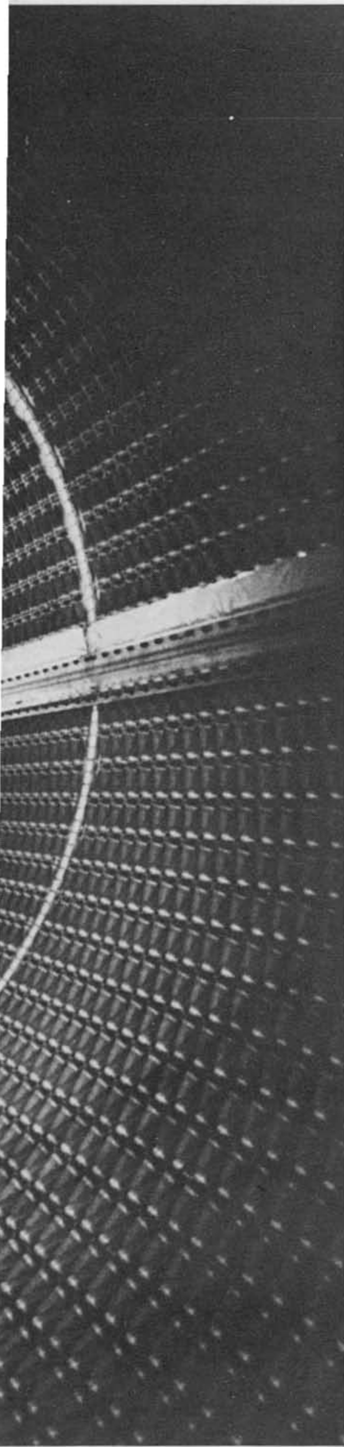


Planche II.5. L'astronaute Edwin E. Aldrin, pilote du module lunaire, marche sur la Lune au cours de la mission Apollo 11. (Avec l'aimable autorisation de la NASA.)

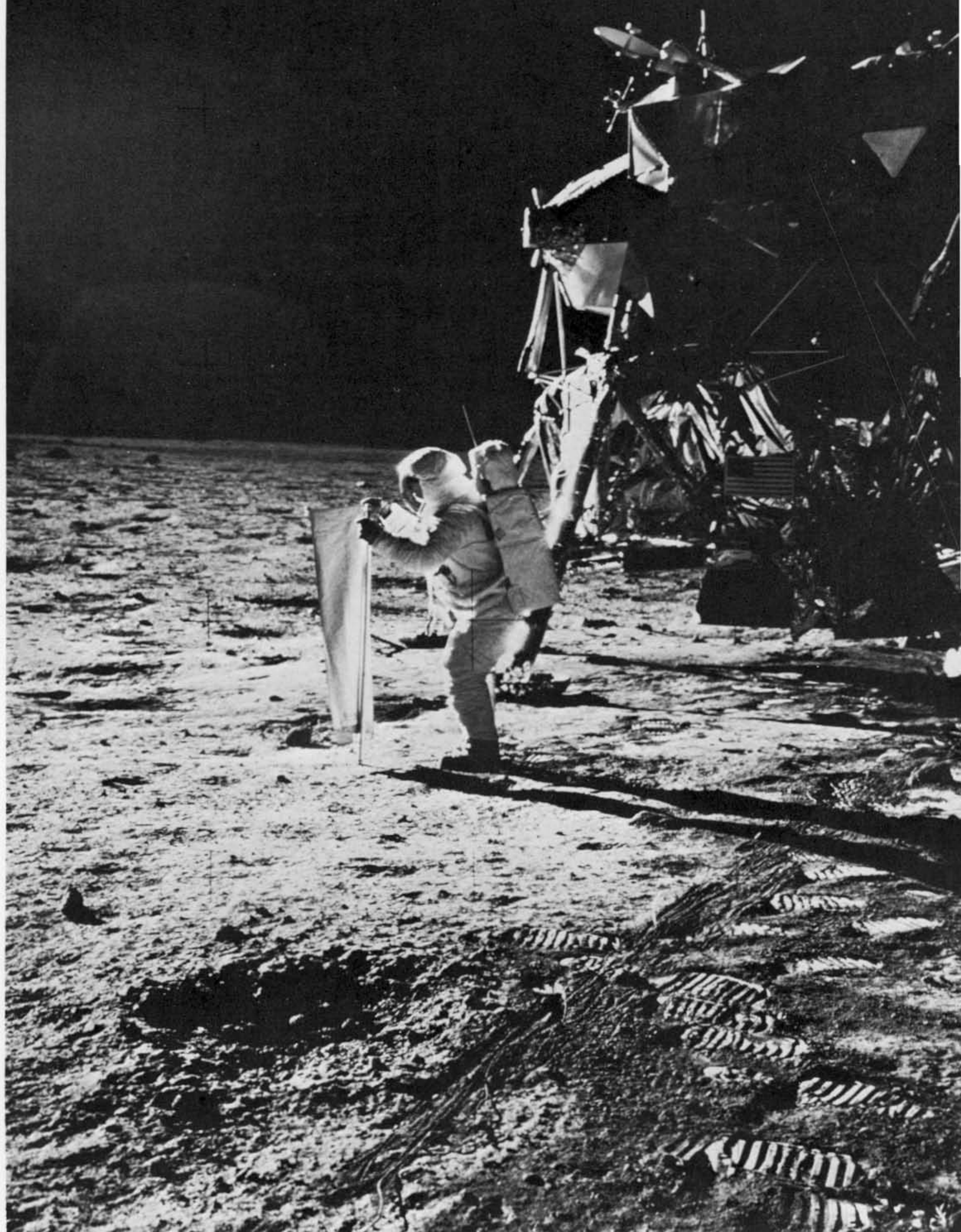
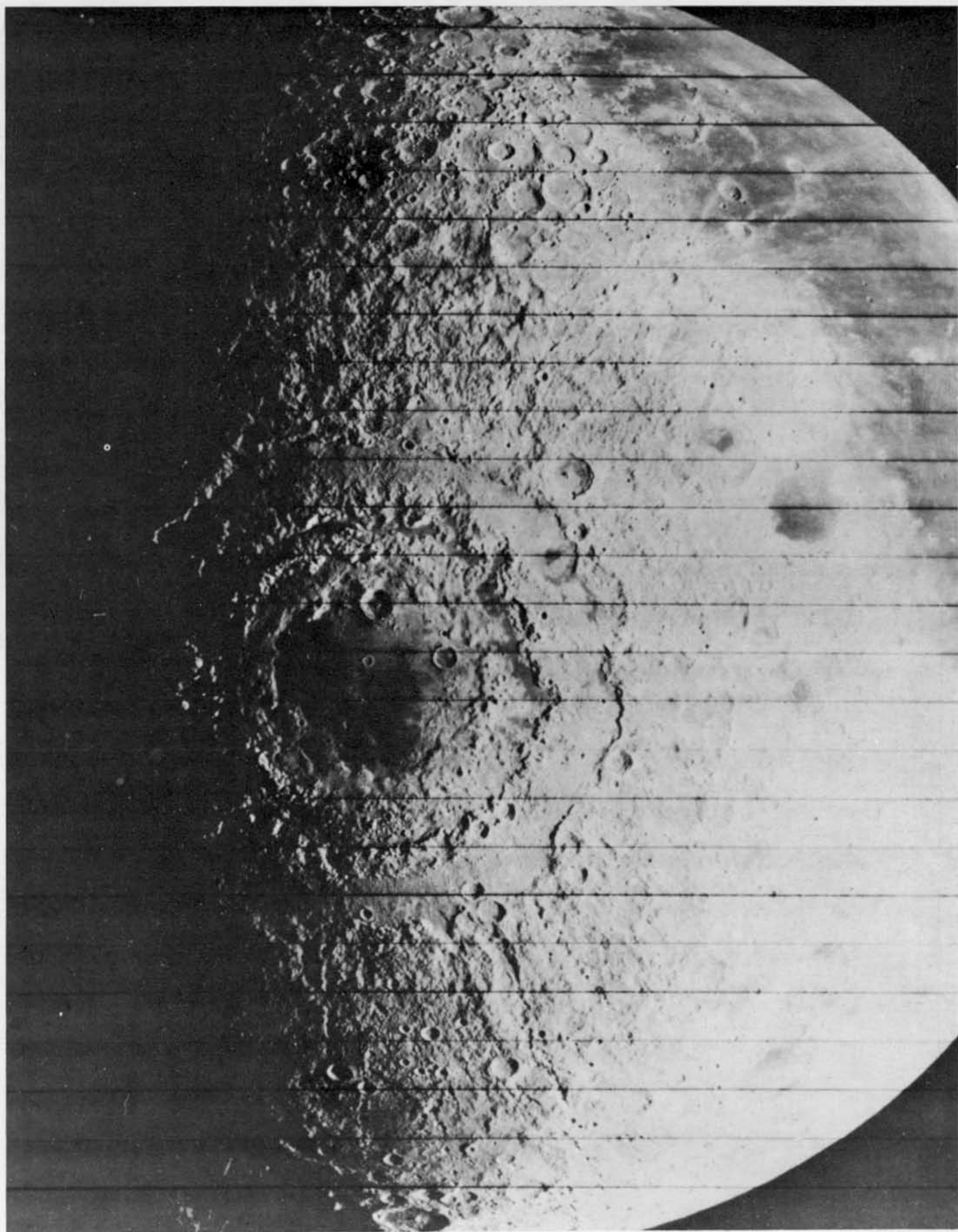


Planche II.6. L'astronaute Edwin E. Aldrin déploie à la surface de la Lune l'appareil de mesure du vent solaire lors de la mission Apollo 11. (Avec l'aimable autorisation de la NASA.)

Planche II.7. Le bassin Oriental photographié à une altitude de 2 700 kilomètres par *Lunar Orbiter IV*. (Avec l'aimable autorisation de la NASA.)



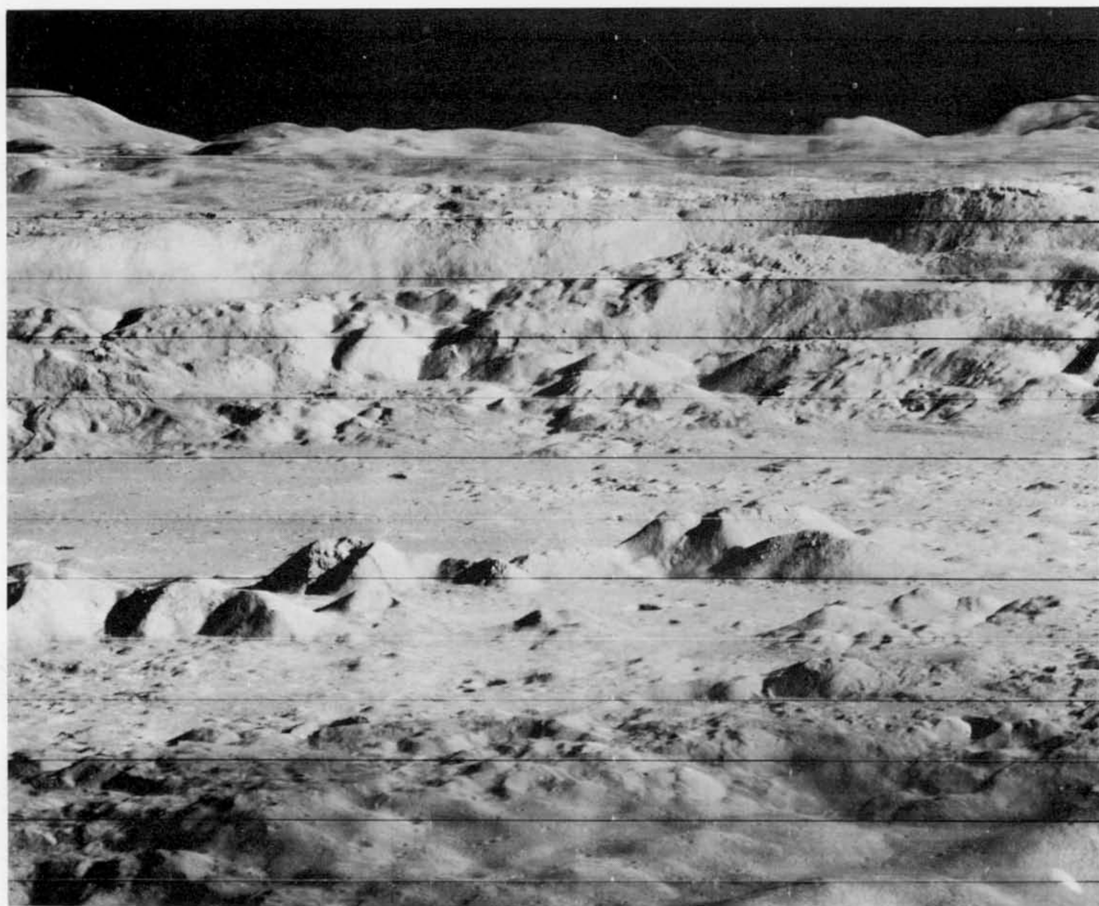


Planche II.8. Photographie du cratère Copernic prise à 45 kilomètres de haut par *Lunar Orbiter II.* (Avec l'aimable autorisation de la NASA.)

Planche II.9. Le vaisseau spatial *Apollo 17,* le 28 août 1972. (Avec l'aimable autorisation de la NASA.)

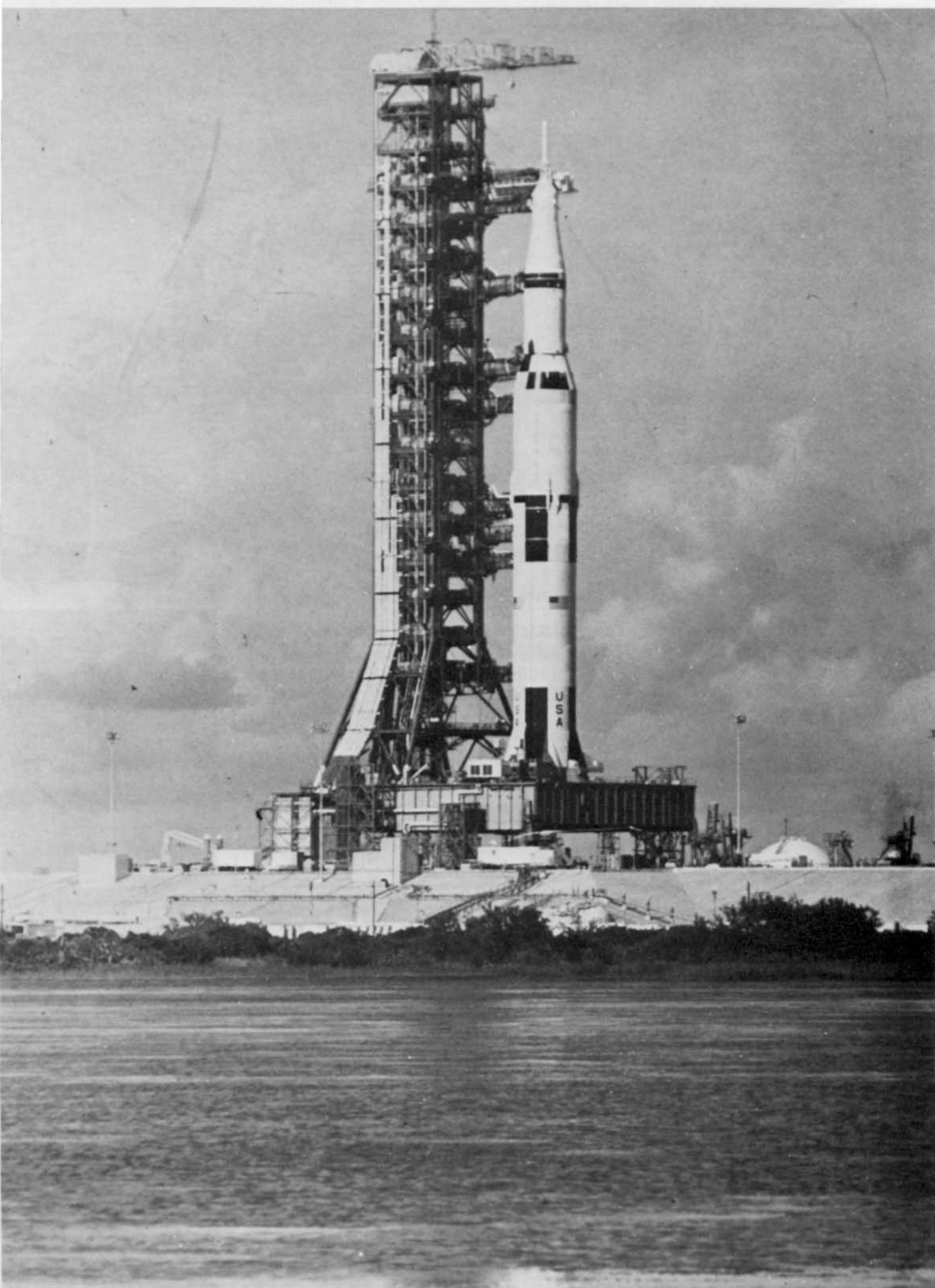


Planche II.10. La surface lunaire. Le cosmonaute Harrison F. Schmitt à côté d'un gros rocher lunaire, durant la mission Apollo 17. Cette scène est un montage de trois photographies. (Avec l'aimable autorisation de la NASA.)



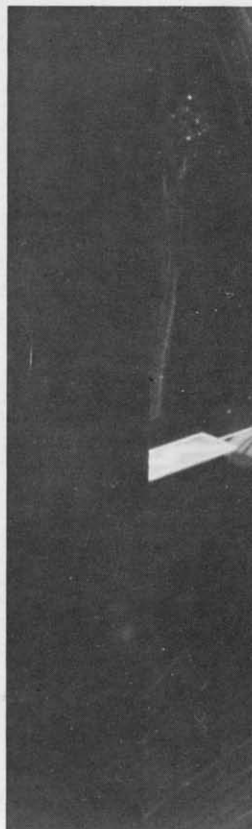


Planche II.11. Maquette d'une future station spatiale. (Avec l'aimable autorisation de la NASA.)









Planche II.12. La Terre, du lever au coucher du Soleil. Cette séquence a été prise par le satellite *ATS-III* à 40 kilomètres environ au-dessus de l'Amérique du Sud, en novembre 1967. Les photos montrent l'ensemble de ce continent, et des portions de l'Amérique du Nord, de l'Afrique, de l'Europe et du Groenland. Des nuages couvrent l'Antarctique. (Avec l'aimable autorisation de la NASA.)





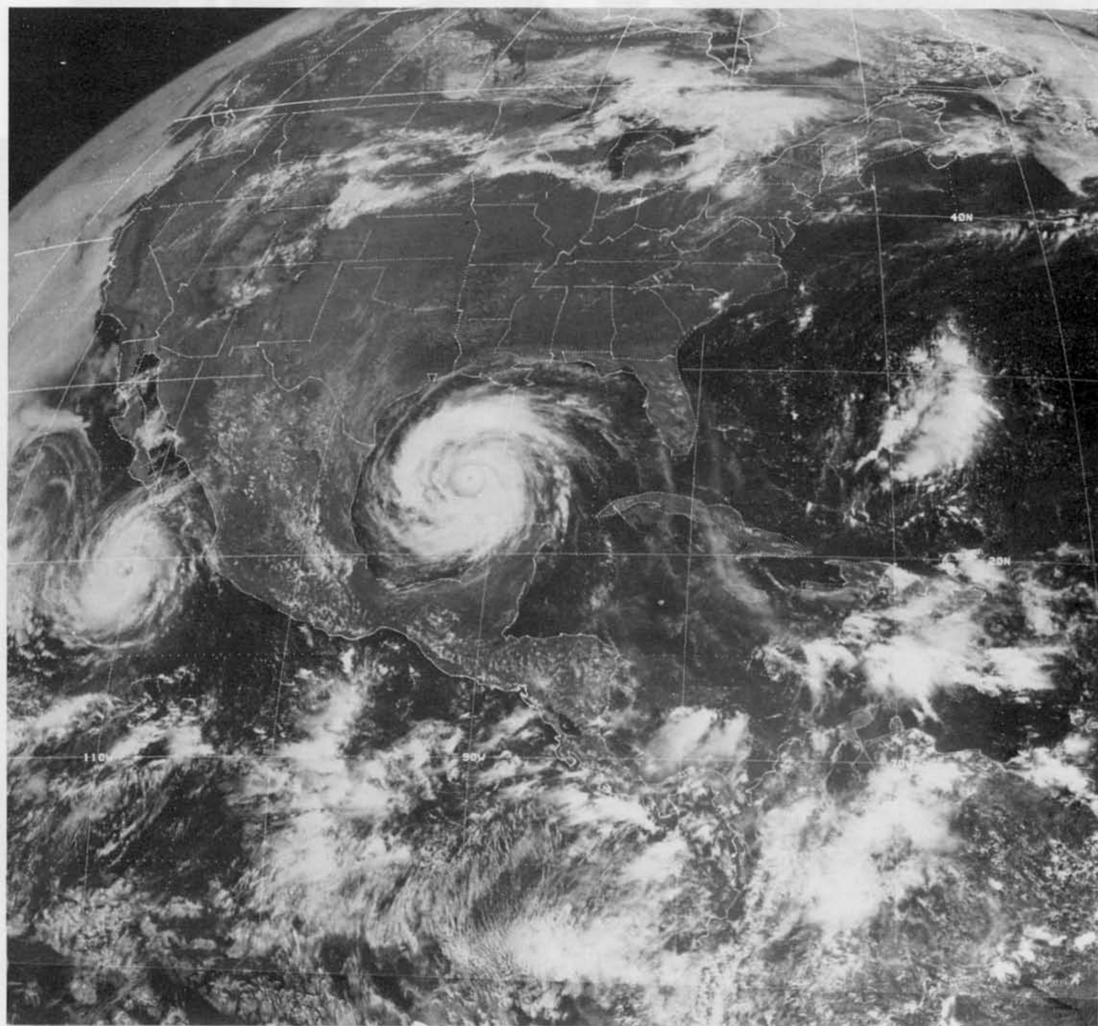


Planche II.13. Photographie météo de la Terre, montrant des tempêtes sur le Pacifique et dans les Caraïbes. (Avec l'aimable autorisation du département du Commerce des Etats-Unis.)









Planche II.14. Sally Ride, la première femme astronaute, se prépare pour le lancement de la navette spatiale *STS 7*, le 18 juin 1983. (Avec l'aimable autorisation de United Press International.)

Planche II.15. La première « marche en liberté dans l'espace », février 1984. L'astronaute Bruce McCandless, sans câble de retenue, à son éloignement maximal de la navette *Challenger*. (Avec l'aimable autorisation de la NASA.)

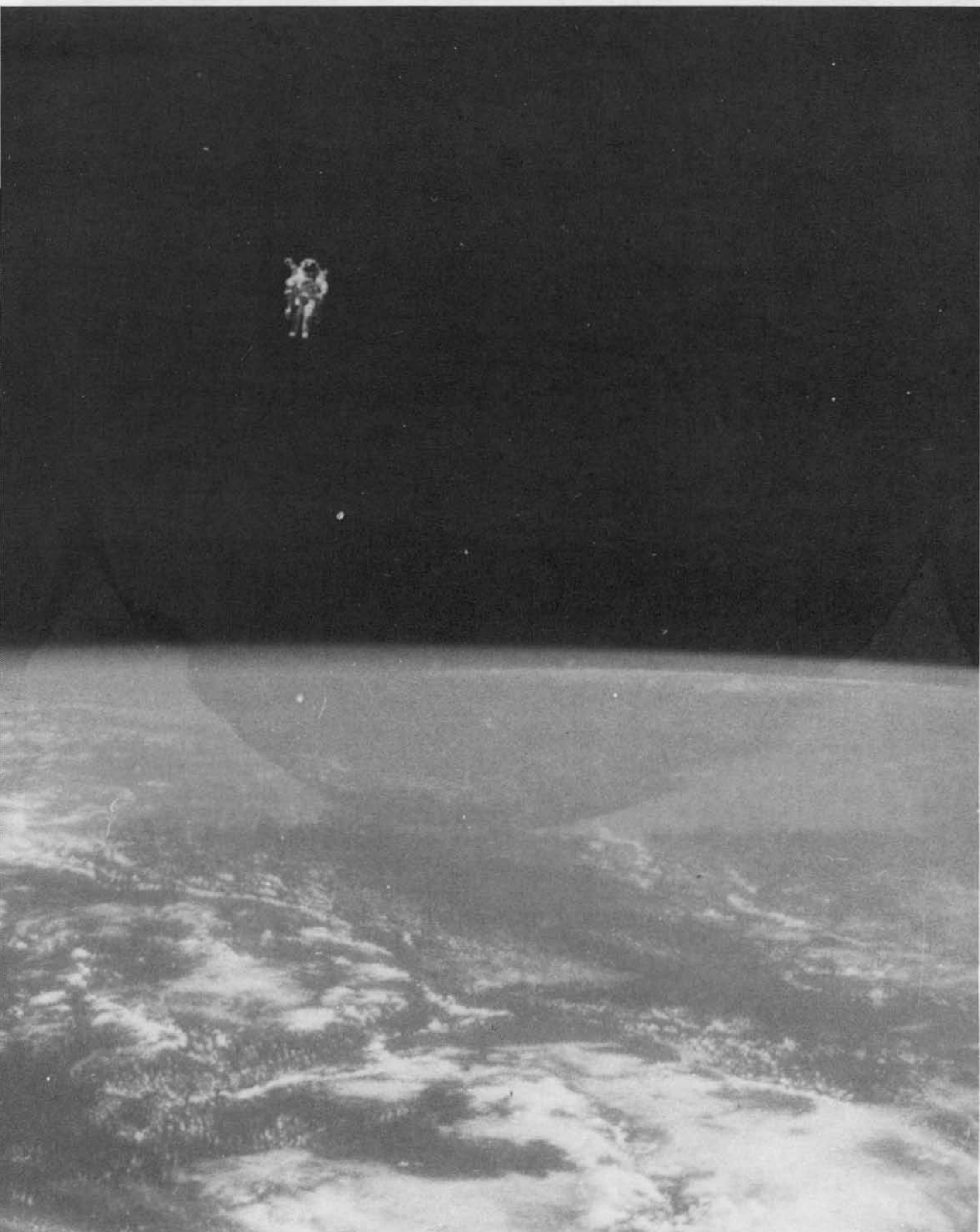




Planche II.16. La Terre, vue d'*Apollo 17,* lors du dernier atterrissage sur la Lune. On voit presque toute l'Afrique et la péninsule arabique. Une épaisse couche de nuages couvre l'Antarctique. (Avec l'aimable autorisation de la NASA.)

que par des méthodes indirectes, en 1952, jusqu'à ce que les sondes jupitériennes, bien sûr, confirment ces résultats à partir de 1973 et nous donnent davantage de détails.

A l'autre bout de l'échelle, une petite planète comme Mars est moins capable que la Terre de retenir même les molécules les plus massives, et son atmosphère est cent fois moins dense que la nôtre. La Lune, avec une vitesse de libération encore plus faible, ne retient rien du tout, et elle est pratiquement dépourvue d'atmosphère.

La température joue aussi un rôle important. D'après la loi de Maxwell-Boltzmann, la vitesse moyenne des particules est proportionnelle à la racine carrée de la température absolue. Si la Terre était à la température de la surface solaire, tous les atomes et les molécules de son atmosphère seraient accélérés quatre ou cinq fois, et elle ne pourrait pas plus retenir l'oxygène et l'azote qu'elle ne retient l'hydrogène et l'hélium.

Ainsi, Mercure a une gravité de surface 2,2 fois plus forte que celle de la Lune, et devrait mieux réussir à garder une atmosphère. Seulement cette petite planète est considérablement plus chaude que la Lune, et se retrouve aussi dépourvue d'atmosphère que notre satellite.

Mars a une gravité de surface à peine supérieure à celle de Mercure, mais sa température est beaucoup plus basse, plus basse même que celle de la Lune. C'est ce qui permet à cette planète d'avoir un peu d'atmosphère, bien plus que la valeur un peu moins faible de sa gravité de surface. Quant aux satellites de Jupiter, ils sont encore plus froids que Mars, mais leur attraction est de l'ordre de celle de la Lune, ce qui ne leur permet pas de retenir une atmosphère. Titan, le gros satellite de Saturne, est à peu près de la même taille, mais si froid qu'il maintient une épaisse atmosphère d'azote. Peut-être en est-il de même pour Triton, le gros satellite de Neptune.

L'ATMOSPHÈRE ORIGINELLE

Le fait que la Terre possède une atmosphère est un argument puissant contre les théories « catastrophiques » de l'origine du système solaire, par exemple la presque-collision entre notre Soleil et une autre étoile. Cela plaide au contraire en faveur de la théorie de la nébuleuse et des planéticules. A mesure que les gaz et les poussières du nuage originel se sont condensés en planéticules, puis que ceux-ci se sont rassemblés en planètes, du gaz a pu se trouver piégé dans cette masse spongieuse, comme l'air dans un tas de neige. La masse s'accroissant, son champ gravitationnel a augmenté et serré les morceaux plus fort les uns contre les autres, ce qui a expulsé les gaz vers la surface. La proportion d'un gaz donné qui allait rester prisonnier de la Terre dépendait de son activité chimique. L'hélium et le néon, sans doute parmi les éléments les plus abondants dans le nuage initial, sont tellement inertes chimiquement qu'ils ont dû ne former aucun composé et s'échapper très vite sous forme de gaz. Aussi, les concentrations d'hélium et de néon sur la Terre ne sont que d'infimes fractions de leurs concentrations dans l'Univers en général. On a pu calculer, par exemple, que la Terre n'avait gardé qu'un atome de néon sur cinquante milliards présents dans le nuage originel, et moins encore – sinon aucun – des atomes d'hélium. Nous disons « aucun », parce que le peu d'hélium présent dans l'atmosphère actuelle peut provenir

entièrement de la fission d'éléments radioactifs, et de gaz piégé dans des cavités souterraines.

L'hydrogène, au contraire, plus léger pourtant que l'hélium ou le néon, a été beaucoup plus facile à garder, parce qu'il s'est combiné avec d'autres substances, en particulier avec l'oxygène pour former de l'eau. On estime ainsi que de l'hydrogène du nuage originel, la Terre a pu garder un atome sur cinq millions.

L'azote et l'oxygène illustrent encore plus nettement cette importance de la réactivité chimique : bien que les molécules de ces deux gaz soient à peu près de même masse, la Terre a gardé un sur six des atomes d'oxygène – très réactif – présents au départ, et seulement un sur 800 000 des atomes d'azote – gaz particulièrement inerte.

Quand nous parlons des gaz de l'atmosphère, nous devons y inclure la vapeur d'eau, et il se pose aussitôt l'intéressante question de l'origine des océans. Au début de l'histoire de la Terre, même si elle n'était que modérément chaude, toute l'eau devait y être sous forme de vapeur. Certains géologues pensent que l'eau était alors entièrement dans l'atmosphère, sous la forme de nuages de vapeur, et qu'elle est tombée en pluie quand la Terre s'est refroidie, formant ainsi les océans. D'autres, en revanche, soutiennent que nos océans se sont surtout formés à partir de l'eau emprisonnée dans la Terre, et peu à peu expulsée vers la surface. Les volcans prouvent qu'il y a encore beaucoup d'eau dans l'écorce terrestre, car les gaz qu'ils émettent sont surtout formés de vapeur d'eau. Dans ce cas, peut-être les océans sont-ils encore en train de grandir, très lentement.

Mais l'atmosphère terrestre a-t-elle toujours été, depuis sa formation, ce qu'elle est maintenant ? Cela semble peu probable. D'abord, l'oxygène moléculaire, qui en forme le cinquième, est un corps si actif que sa présence sous forme libre, non combinée, est extrêmement improbable, à moins que l'on en produise en permanence. D'autre part, aucune autre planète n'a une atmosphère comme la nôtre, ce qui incite fortement à penser que l'atmosphère terrestre résulte d'événements exceptionnels (par exemple, la présence de la vie sur cette planète-ci et pas sur les autres).

Selon Harold Urey, de nombreuses raisons permettent de penser que l'atmosphère originelle était formée surtout d'ammoniac et de méthane. Dans l'Univers, les éléments les plus répandus sont l'hydrogène (de loin), puis l'hélium, le carbone, l'azote et l'oxygène. En présence de cette surabondance d'hydrogène, le carbone donnerait du méthane (CH_4), l'azote de l'ammoniac (NH_3) et l'oxygène de l'eau (H_2O). L'hélium et le surplus d'hydrogène s'échappent, bien entendu. L'eau forme les océans. Le méthane et l'ammoniac, gaz relativement lourds, sont retenus par la gravitation et forment l'essentiel de l'atmosphère terrestre.

Si toutes les planètes suffisamment massives pour garder une atmosphère ont effectivement débuté avec une atmosphère de ce type, elles ne pouvaient pas toutes la garder telle quelle. Le rayonnement ultraviolet du Soleil devait y apporter des changements. Ces changements sont négligeables pour les planètes extérieures, qui reçoivent relativement peu de rayonnement d'un Soleil très lointain, et qui ont d'autre part des atmosphères très vastes, capables d'absorber beaucoup de rayonnement sans changements substantiels. On s'attend donc à ce que ces planètes aient gardé jusqu'ici des atmosphères formées d'hydrogène, d'hélium, de méthane et d'ammoniac.

Mais il en va tout autrement pour les planètes intérieures, Mars, la Terre, Vénus et Mercure, et pour la Lune. Parmi celles-ci, Mercure et la Lune sont trop petites, trop chaudes (ou les deux à la fois) pour pouvoir garder une atmosphère. Restent Mars, la Terre et Vénus, avec des atmosphères assez ténues, composées au départ d'ammoniac, de méthane et d'eau. Que va-t-il se passer ?

L'ultraviolet qui atteint les molécules d'eau présentes dans la haute atmosphère va les casser en donnant de l'hydrogène et de l'oxygène (*photodissociation*). L'hydrogène s'échappe, laissant derrière lui l'oxygène. Mais les molécules de celui-ci vont réagir avec la quasi-totalité des molécules présentes dans leur voisinage. Avec le méthane, elles vont donner du gaz carbonique (CO_2) et de l'eau (H_2O). Avec l'ammoniac (NH_3) elles vont former de l'azote (N_2) et de l'eau. Lentement mais sûrement, l'atmosphère de méthane et d'ammoniac devient une atmosphère de gaz carbonique et d'azote. Celui-ci réagit lentement avec les minéraux de l'écorce terrestre pour former des nitrates, ce qui laisse le gaz carbonique comme constituant principal de l'atmosphère.

Cette évolution va-t-elle se poursuivre ? L'eau va-t-elle continuer à se dissocier, l'hydrogène à s'évader dans l'espace, et l'oxygène à s'amasser dans l'atmosphère ? Et si l'oxygène s'accumule et ne trouve autour de lui que du gaz carbonique, avec lequel il ne peut pas se combiner, sa concentration ne va-t-elle pas augmenter, devenir plus grande que celle du gaz carbonique, et rendre compte de l'atmosphère actuelle ? Absolument pas !

Dès que le gaz carbonique est devenu le constituant principal de l'atmosphère, le rayonnement ultraviolet cesse de provoquer des changements par dissociation de l'eau. Quand l'oxygène libre commence à se concentrer, une mince couche d'ozone se forme dans la haute atmosphère. Cette couche absorbe l'ultraviolet et l'empêche d'atteindre les couches plus basses (et donc d'y provoquer la photodissociation) : l'atmosphère de gaz carbonique est stable.

Mais le gaz carbonique introduit l'effet de serre (voir chapitre 4). Si l'atmosphère de gaz carbonique est peu dense et loin du Soleil, et si elle contient peu d'eau de toute façon, comme dans le cas de Mars, l'effet est faible.

Mais imaginons une planète comme la Terre, et aussi rapprochée (ou même plus proche) du Soleil. L'effet de serre va être considérable. La température va s'élever, vaporisant davantage les océans. La vapeur d'eau produite va augmenter l'effet de serre – ainsi d'ailleurs que le gaz carbonique supplémentaire libéré par l'échauffement de l'écorce terrestre. Finalement, la planète va être torride, toute son eau va se trouver sous forme de vapeur dans une atmosphère épaisse et chargée de nuages masquant éternellement la surface, cette atmosphère étant formée essentiellement de gaz carbonique.

C'est le portrait de Vénus, qui a dû effectivement subir un effet de serre divergent de ce type, déclenché par la chaleur supplémentaire reçue, par rapport à la Terre, à cause de sa proximité plus grande du Soleil.

La Terre n'a évolué ni dans le sens de Mars, ni dans celui de Vénus. L'azote de son atmosphère ne s'est pas entièrement englouti dans l'écorce, laissant une atmosphère ténue de gaz carbonique froid, comme sur Mars. Et l'effet de serre n'en a pas non plus fait un désert brûlant comme sur Vénus. Quelque chose d'autre s'est produit, l'apparition de la vie (peut-être alors que l'atmosphère était encore au stade ammoniac-méthane).

Les réactions chimiques provoquées dans l'Océan par les êtres vivants ont dissocié les composés azotés, libéré l'azote et maintenu ainsi une concentration importante de ce gaz dans l'atmosphère. De plus, les cellules vivantes ont acquis la capacité de dissocier l'eau en hydrogène et oxygène en utilisant le rayonnement *visible*, qui n'est pas arrêté par la couche d'ozone. L'hydrogène s'est combiné à l'oxyde de carbone pour former les molécules compliquées qui constituent les cellules, tandis que l'oxygène libéré s'est répandu dans l'atmosphère. Ainsi, grâce à la vie, l'atmosphère d'azote et de gaz carbonique est devenue une atmosphère d'azote et d'oxygène. L'effet de serre a été très limité. La Terre est restée fraîche et capable de garder ses deux trésors uniques : un Océan d'eau liquide et une atmosphère riche en oxygène libre.

D'ailleurs, notre atmosphère oxygénée caractérise peut-être seulement le dernier dixième de l'existence de la Terre. Il est possible que cette atmosphère, il y a seulement 600 millions d'années, ait contenu dix fois moins d'oxygène qu'à présent.

Mais enfin, cet oxygène, nous l'avons maintenant : soyons donc reconnaissants à la vie, qui a rendu possible l'oxygène atmosphérique, et à cet oxygène qui rend à son tour possible notre vie.

Chapitre 6

Les éléments

Le tableau périodique

Nous ne nous sommes encore occupés que des corps volumineux qui se présentent dans l'Univers – les étoiles, les galaxies, le système solaire et la Terre avec son atmosphère. Etudions maintenant la nature des substances qui entrent dans la composition de tous ces corps.

LES PREMIÈRES THÉORIES

Les philosophes grecs (qui ont envisagé la plupart des problèmes de façon théorique et spéculative) avaient décidé que la Terre était faite de très peu *d'éléments* ou substances de base. Vers 430 av. J.-C., Empédocle fixait leur nombre à quatre – la terre, l'air, l'eau et le feu. Un siècle plus tard, Aristote supposait que le ciel était constitué par un cinquième élément, *l'éther*. Les alchimistes du Moyen Age qui, à la suite des Grecs, ont étudié la matière, quoique empêtrés dans la magie et le charlatanisme, sont arrivés à des conclusions plus judicieuses et raisonnables car, au moins, ils manipulaient les matériaux à propos desquels ils spéculaient.

Tentant d'expliquer la diversité des propriétés des substances, les alchimistes ont associé ces propriétés à des éléments de contrôle qu'ils ajoutèrent à la liste. Pour eux, le mercure était l'élément qui dotait les substances de propriétés métalliques, tandis que le soufre leur permettait d'être combustibles. L'un des derniers et des plus illustres des alchimistes, le médecin suisse du XVI[e] siècle Theophrastus Bombastus von Hohenheim,

plus connu sous le nom de Paracelse, ajouta le sel comme élément de résistance à la chaleur.

Selon le raisonnement des alchimistes, il suffisait d'ajouter ou de retrancher des éléments en quantités convenables pour qu'une substance se change en une autre. Un métal comme le plomb, par exemple, pouvait être changé en or par addition de la bonne quantité de mercure. La quête du processus de conversion d'un métal vil en or s'est poursuivie pendant des siècles. En cours de route les alchimistes ont découvert des substances beaucoup plus importantes que l'or – comme les acides minéraux et le phosphore.

Les acides minéraux – l'acide nitrique, l'acide chlorhydrique, et surtout l'acide sulfurique (préparé pour la première fois vers l'an 1300) – ont pratiquement marqué le début d'une révolution dans les expériences des alchimistes. Ces substances étaient des acides beaucoup plus forts que ceux connus auparavant (dont l'acide acétique du vinaigre, le plus fort de tous) ; ils permettaient de décomposer les substances sans chauffer et sans attendre trop longtemps. Encore à présent, les acides minéraux, en particulier l'acide sulfurique, sont d'une importance vitale pour l'industrie. On dit que le degré d'industrialisation d'un pays peut se mesurer à sa consommation annuelle d'acide sulfurique.

Néanmoins, peu d'alchimistes se sont laissé distraire par ces voies annexes de ce qu'ils considéraient être leur principal sujet de recherche. Des membres peu scrupuleux de la corporation se livraient purement et simplement à des trucages, produisant de l'or par prestidigitation pour soutirer à de riches mécènes ce que nous appellerions maintenant des « crédits de recherche ». La profession en acquit si mauvaise réputation qu'il fallut abandonner le titre même d'*alchimiste.* Au XVII[e] siècle, l'*alchimiste* était devenu *chimiste* et l'*alchimie* s'était élevée à la dignité d'une science baptisée *chimie.*

A l'aurore resplendissante de cette science, Robert Boyle, l'auteur de la loi des gaz parfaits (voir le chapitre 5), fut l'un des premiers de ces nouveaux chimistes. Dans son ouvrage *The Sceptical Chymist* (le chimiste sceptique), publié en 1661, Boyle a formulé le premier la définition spécifique moderne d'un élément : une substance de base qui peut se combiner avec d'autres éléments pour former des *composés* et qui, inversement, après avoir été extraite d'un composé, ne peut être réduite à une substance plus simple.

Toutefois, Boyle avait encore un point de vue moyenâgeux sur ce qui appartenait à la catégorie des éléments. Il croyait par exemple que l'or n'était pas un élément et pouvait être formé d'une façon ou d'une autre à partir d'autres métaux. Il en était de même, à vrai dire, de son contemporain Isaac Newton, qui consacra beaucoup de temps à l'alchimie. (En fait, François-Joseph, empereur d'Autriche-Hongrie, subventionnait encore des expériences de fabrication de l'or en 1867.)

Un siècle après Boyle, les travaux chimiques expérimentaux ont permis d'établir sans ambiguïté quelles étaient les substances qui pouvaient être décomposées en substances plus simples et celles pour lesquelles c'était impossible. Henry Cavendish a montré que l'hydrogène pouvait se combiner avec l'oxygène pour former de l'eau, de sorte que l'eau ne pouvait être un élément. Plus tard, Lavoisier a décomposé le prétendu élément nommé air en oxygène et azote. Il devint patent qu'aucun des éléments des Grecs n'était un élément selon le critère de Boyle.

Pour ce qui est des éléments des alchimistes, le mercure et le soufre se révélèrent être des éléments « au sens de Boyle ». Mais ce fut aussi le cas du fer, de l'étain, du plomb, du cuivre, de l'argent, de l'or, et de substances non métalliques comme le phosphore, le carbone et l'arsenic. L'« élément » sel de Paracelse fut finalement décomposé en deux substances plus simples.

Bien entendu, la définition des éléments était dépendante de la chimie de l'époque. Tant qu'on ne pouvait la décomposer par les techniques chimiques connues, une substance pouvait toujours être considérée comme un élément. La liste de trente-trois éléments de Lavoisier comprenait par exemple des substances comme la chaux et la magnésie. Mais quatorze ans après la fin de Lavoisier sous la guillotine révolutionnaire, le chimiste anglais Humphry Davy décomposa, au moyen d'un courant électrique, la chaux en oxygène et en un nouvel élément qu'il baptisa *calcium,* et il sépara de la même façon la magnésie en oxygène et en un autre nouvel élément qu'il appela *magnésium.*

En revanche, Davy put montrer qu'un gaz vert que le chimiste suédois Carl Wilhelm Scheele avait préparé à partir de l'acide chlorhydrique n'était nullement un composé d'acide chlorhydrique et d'oxygène, comme on le pensait, mais un authentique élément qu'il désigna sous le nom de *chlore* (d'après le mot grec qui signifie « vert »).

LA THÉORIE ATOMIQUE

Au début du XIXe siècle s'est répandu un point de vue radicalement nouveau sur les éléments, qui remontait en fait à quelques penseurs grecs de l'Antiquité ; ceux-ci ont finalement contribué à ce qui a été peut-être le concept le plus important dans la compréhension de la matière.

Les Grecs se demandaient si la matière était continue ou discontinue, c'est-à-dire si on pouvait la diviser et la subdiviser indéfiniment en une poussière toujours plus fine, ou si elle s'avérerait finalement être constituée de particules indivisibles. Vers 450 av. J.-C., Leucippe et son élève Démocrite prétendaient que la deuxième hypothèse était la bonne. Démocrite donna un nom à ces particules : il les appela *atomes* (ce qui signifie « indivisibles »). Il suggéra même que des substances différentes étaient composées d'atomes, ou de combinaisons d'atomes différents, et qu'une substance pouvait être convertie en une autre par réarrangement des atomes. Considérant que tout cela n'était après tout que conjectures raisonnables, on est frappé par l'exactitude de ses intuitions. Cette idée, qui peut maintenant nous sembler évidente, l'était en fait si peu à l'époque que Platon et Aristote la rejetèrent immédiatement.

Elle se maintint malgré tout dans l'enseignement d'Epicure qui écrivait vers 300 av. J.-C., et dans l'école philosophique (l'épicurisme) à laquelle il donna naissance. Un illustre épicurien, le philosophe romain Lucrèce, incorpora les idées atomiques dans un long poème, *De la nature,* vers 60 av. J.-C. Une copie délabrée survécut jusqu'au Moyen Age, et le poème de Lucrèce fut une des premières œuvres à bénéficier de l'invention de l'imprimerie.

Ainsi, la notion d'atome n'a jamais disparu totalement du subconscient des érudits occidentaux. Parmi les atomistes ressortent particulièrement, à l'aurore de la science, le philosophe italien Giordano Bruno et le

philosophe français Pierre Gassendi. Bruno professait de nombreux points de vue scientifiques non orthodoxes, comme la croyance en un Univers infini dont les étoiles étaient des soleils lointains autour desquels tournaient des planètes, et il s'exprimait sans retenue. Il fut brûlé comme hérétique en 1600 ; c'est le plus éminent des martyrs de la révolution scientifique. Les Soviétiques ont donné son nom à un cratère de la face cachée de la Lune.

Les opinions de Gassendi ont influencé Boyle dont les propres expériences, montrant qu'il était facile de comprimer ou dilater les gaz, semblaient indiquer que ceux-ci devaient être composés de particules très espacées. Boyle et Newton figuraient donc parmi les atomistes convaincus du XVII[e] siècle.

En 1799, le chimiste français Joseph Louis Proust montra que le carbonate de cuivre contenait des proportions bien définies en poids de cuivre, de carbone et d'oxygène, quelle qu'en soit la préparation. Ces proportions étaient représentées par des nombres entiers petits : 5 à 4 à 1. Il poursuivit en montrant qu'il en allait de façon analogue pour quantités d'autres composés.

Cet état de choses pouvait s'expliquer en supposant que les composés étaient constitués de groupements de quelques fragments de chaque élément, qui ne pouvaient se combiner qu'en gardant leur intégrité. C'est ce que remarqua le chimiste anglais John Dalton en 1803 ; il publia en 1808 un ouvrage dans lequel il établissait l'unité des connaissances acquises en chimie au cours des cent cinquante dernières années, si l'on supposait que toute matière était constituée d'atomes indivisibles. (Dalton conserva le vieux mot grec, en hommage aux penseurs de l'Antiquité.) Sa *théorie atomique* convainquit bien vite la plupart des chimistes.

Selon Dalton, chaque élément possède son type propre d'atomes, et toute quantité de cet élément est faite d'atomes identiques de ce type. C'est la nature de ses atomes qui distingue un élément d'un autre. La différence physique entre atomes la plus évidente est leur poids. Ainsi, les atomes de soufre sont plus lourds que les atomes d'oxygène, eux-mêmes plus lourds que les atomes d'azote, qui sont plus lourds que les atomes de carbone, à leur tour plus lourds que les atomes d'hydrogène.

En appliquant la théorie atomique, le chimiste italien Amedeo Avogadro a montré que l'on avait intérêt à supposer que des volumes de gaz égaux (quelle que soit leur nature) renfermaient des nombres de particules égaux. C'est *l'hypothèse d'Avogadro.* On supposa d'abord que ces particules étaient des atomes, mais elles se sont révélées finalement composées, dans la plupart des cas, de petits groupements d'atomes appelés *molécules.* Lorsqu'une molécule contient des atomes de divers types (comme la molécule d'eau, qui est constituée d'un atome d'oxygène et de deux atomes d'hydrogène), c'est la molécule d'un *composé chimique.*

Il devint évidemment important de mesurer les poids relatifs des divers atomes pour déterminer les *poids atomiques* des éléments. Il était bien entendu hors de question de peser les atomes proprement dits, avec les techniques disponibles au XIX[e] siècle. Mais en pesant la quantité de chaque élément isolé d'un composé, et en connaissant le comportement chimique de l'élément, on pouvait déduire les poids relatifs des atomes. C'est le chimiste suédois Jöns Jacob Berzelius qui fit le premier cette étude systématique. Il publia en 1828 une liste des poids atomiques fondée sur deux références : il adoptait arbitrairement comme poids atomiques 100 pour l'oxygène et 1 pour l'hydrogène.

Le système de Berzelius ne fut pas immédiatement accepté ; mais en 1860, au premier Congrès international de chimie, à Karlsruhe, le chimiste italien Stanislao Cannizzaro présentait de nouvelles méthodes pour déterminer les poids atomiques, utilisant l'hypothèse d'Avogadro, jusqu'alors négligée. Cannizzaro exposa si énergiquement son point de vue que la communauté des chimistes fut convaincue.

On adopta à cette époque comme référence le poids de l'oxygène plutôt que celui de l'hydrogène, car l'oxygène est plus enclin à se combiner avec d'autres éléments (et c'est cette combinaison avec d'autres éléments qui constituait l'étape clé dans la méthode ordinaire de détermination des poids atomiques). Le poids atomique de l'oxygène fut arbitrairement fixé par le chimiste belge Jean Servais Stas, en 1850, à la valeur 16 exactement, de sorte que le poids atomique de l'hydrogène, le plus léger des éléments connus, se trouvait être voisin de 1 – 1,0080 pour être exact.

Sans cesse, depuis Cannizzaro, les chimistes se sont efforcés de déterminer les poids atomiques avec une précision toujours meilleure.

Ce travail a culminé, tout au moins en ce qui concerne les méthodes purement chimiques, avec le chimiste américain Theodore William Richards qui, à partir de 1904, a déterminé les poids atomiques avec une précision sans précédent. Cela lui a valu le prix Nobel de chimie en 1914. A la suite des découvertes ultérieures sur la constitution physique des atomes, les chiffres donnés par Richards ont été encore améliorés.

Tout au long du XIXe siècle, en dépit des travaux effectués sur les atomes et les molécules, et bien que les chercheurs aient été en général convaincus de leur réalité, il n'existait aucune preuve directe qu'ils pouvaient être plus que des représentations abstraites commodes. Quelques éminents scientifiques, comme le chimiste allemand Wilhelm Ostwald, refusaient d'y voir autre chose. Pour eux, ces particules étaient pratiques mais pas « réelles ».

C'est le *mouvement brownien* qui a établi la réalité des molécules. Ce phénomène fut observé pour la première fois en 1827 par le botaniste écossais Robert Brown, qui remarqua que des grains de pollen en suspension dans l'eau s'agitaient de façon désordonnée. On pensa tout d'abord que cette agitation était due à la nature vivante des grains de pollen, mais des particules de teinture aussi petites et parfaitement inanimées montrèrent la même agitation.

En 1863, on suggéra pour la première fois que le mouvement était dû à un bombardement inégal des particules de la part des molécules d'eau environnantes. Pour de gros objets, un léger déséquilibre du nombre de molécules frappant à gauche et à droite devait être sans conséquence. Pour des objets microscopiques, soumis au bombardement de peut-être seulement quelques centaines de molécules par seconde, un déséquilibre de quelques unités – d'un côté ou de l'autre – pouvait causer une agitation perceptible. Le mouvement aléatoire des minuscules particules est presque une preuve visible de la *discontinuité* de l'eau, et de la matière en général.

C'est Einstein qui a réalisé l'analyse théorique de cette représentation du mouvement brownien, et montré comment on pouvait déduire la taille des molécules d'eau de l'amplitude des petits mouvements des particules de teinture. En 1908, le physicien français Jean Perrin a étudié la manière dont les particules se répartissent verticalement dans l'eau sous l'influence de la gravité. Au tassement vers le bas de cette distribution s'opposent les collisions des molécules arrivant par en dessous, de sorte qu'un

mouvement brownien s'oppose à l'attraction gravitationnelle. Perrin utilisa cette découverte pour calculer la taille des molécules d'eau en se servant de l'équation trouvée par Einstein, et même Ostwald dut se rendre à l'évidence. Perrin reçut le prix Nobel de physique pour ses travaux en 1926.

Ainsi, les atomes n'ont cessé d'évoluer du statut d'abstractions semi-mystiques vers celui d'objets presque tangibles. En fait, on peut dire aujourd'hui que l'on est enfin parvenu à « voir » l'atome. Cela a été réalisé au moyen du *microscope ionique à effet de champ,* inventé par Erwin W. Mueller de l'université de Pennsylvanie en 1955. Son système arrache des ions positifs de la pointe d'une aiguille très fine pour les projeter sur un écran fluorescent, de façon à produire une image de la pointe de l'aiguille agrandie cinq millions de fois. L'image révèle réellement les atomes individuels de l'aiguille sous forme de petits points brillants. Cette technique a pu être améliorée au point d'obtenir des images d'atomes isolés. Le physicien américain Albert Victor Crewe a annoncé en 1970 la détection d'atomes individuels d'uranium et de thorium au moyen d'un microscope électronique à balayage.

LE TABLEAU PÉRIODIQUE DE MENDELEÏEV

A mesure que la liste des éléments s'allongeait, les chimistes du XIXe siècle avaient le sentiment de s'empêtrer dans une jungle de plus en plus épaisse. Chaque élément avait des propriétés différentes, et ils ne pouvaient distinguer aucun ordre sous-jacent à cette liste. L'essence même de la science consistant à essayer de discerner un ordre dans le désordre apparent, les scientifiques se mirent à la recherche d'une certaine régularité dans les propriétés des éléments.

En 1862, lorsque Cannizzaro eut établi le rôle du poids atomique parmi les outils importants de la chimie, un géologue français, Alexandre Émile Béguyer de Chancourtois, s'aperçut qu'il pouvait disposer les éléments par ordre de poids atomique croissant, sous forme de tableau, de sorte que les éléments doués de propriétés similaires tombaient dans les mêmes colonnes. Deux ans plus tard, un chimiste britannique, John Alexander Reina Newlands, parvint indépendamment à la même disposition. Mais ces deux chercheurs furent ignorés ou ridiculisés. Aucun ne réussit alors à publier ses propositions de façon sérieuse. Bien des années plus tard, lorsque l'importance du tableau périodique fut universellement reconnue, leurs articles furent enfin publiés. Newlands reçut même une médaille.

C'est au chimiste russe Dimitri Ivanovitch Mendeleïev qu'est revenu finalement l'honneur d'ordonner la jungle des éléments. En 1869, Mendeleïev et le chimiste allemand Julius Lothar Meyer ont proposé des tables des éléments qui faisaient ressortir essentiellement les mêmes propriétés que Chancourtois et Newlands. Mais Mendeleïev a eu le mérite et le courage de pousser cette idée plus loin que ses prédécesseurs ne l'avaient fait.

Tout d'abord, le *tableau périodique* de Mendeleïev (ainsi nommé car s'y manifestait la réapparition périodique de propriétés chimiques similaires) était plus compliqué que celui de Newlands et plus proche de ce que l'on estime maintenant correct (voir le tableau 6.1). Deuxièmement, lorsque les propriétés d'un élément n'étaient pas en accord avec son emplacement

selon l'ordre des masses atomiques, Mendeleïev altérait cet ordre sans sourciller, estimant que les propriétés étaient plus importantes que le poids atomique. L'avenir devait lui donner raison, comme nous le verrons plus loin dans ce chapitre. Par exemple le tellurium, avec son poids atomique de 127,61, aurait dû venir juste après l'iode, dont le poids est de 126,91. Mais dans le tableau, lorsqu'on met le tellurium avant l'iode, il se retrouve sous le sélénium auquel il ressemble étroitement, et l'iode vient sous son cousin le brome.

Finalement, et encore plus important, chaque fois que Mendeleïev n'arrivait pas à trouver une autre façon de faire fonctionner son arrangement, il n'hésitait pas à laisser des cases blanches dans le tableau et à proclamer vaillamment qu'il fallait découvrir les éléments qui appartenaient à ces cases. Il est même allé plus loin. Pour trois des cases vides il donna une description des éléments qui devaient remplir chacune d'elles en se guidant d'après les éléments qui se trouvaient au-dessus et en dessous dans le tableau. C'est là que Mendeleïev eut un coup de chance. Chacun des trois éléments qu'il avait prédits a été découvert de son vivant de sorte qu'il a pu assister au triomphe de son système. En 1875, le chimiste français Lecoq de Boisbaudran découvrit le premier de ces trois éléments manquants et le baptisa *gallium* (d'après l'ancien nom de la France en latin). En 1879, le chimiste suédois Lars Fredrik Nilson trouva le second et le nomma *scandium* (d'après Scandinavie). Et en 1886, le chimiste allemand Clemens Alexander Winkler isola le troisième qu'il appela *germanium* (d'après Germanie). Ces trois éléments avaient bien les propriétés annoncées par Mendeleïev.

LES NUMÉROS ATOMIQUES

La découverte des rayons X par Roentgen amorça une nouvelle étape dans l'histoire du tableau périodique. En 1911, le physicien britannique Charles Glover Barkla découvrit que les rayons X émis par un métal avaient un pouvoir de pénétration bien défini qui dépendait de ce métal ; en d'autres termes, chaque élément produit ses propres *rayons X caractéristiques.* Barkla s'est vu décerner le prix Nobel de physique en 1917 pour cette découverte.

On se demandait si les rayons X étaient constitués par un courant de petites particules ou étaient plutôt des ondes à la façon de la lumière. Une façon de s'en assurer consistait à vérifier si les rayons X pouvaient être *diffractés* (c'est-à-dire obligés à changer de direction) par un *réseau de diffraction* constitué par une série de fines rayures. Mais pour que le phénomène de diffraction se produise réellement, la distance entre les rayures devait être de l'ordre de la longueur d'onde du rayonnement. Les rayures les plus serrées réalisables suffisaient pour la lumière ordinaire, mais le pouvoir de pénétration des X laissait à penser que, dans la mesure où il s'agissait effectivement d'ondes, celles-ci seraient de longueur bien inférieure à la longueur d'onde de la lumière. Ainsi, aucun réseau de diffraction ordinaire ne permettrait de diffracter les rayons X.

Mais il vint à l'esprit du physicien allemand Max Theodore Felix von Laue que les cristaux constituaient des réseaux de diffraction naturels beaucoup plus fins que n'importe quel réseau artificiel. Un *cristal* est un solide dont la forme géométrique est simple ; ses faces planes se joignent

1 Hydrogène (H) 1,008								
3 Lithium (Li) 6,939	**4** Béryllium (Be) 9,012							
11 Sodium (Na) 22,990	**12** Magnésium (Mg) 24,312							
19 Potassium (K) 39,102	**20** Calcium (Ca) 40,08	**21** Scandium (Sc) 44,956	**22** Titane (Ti) 47,90	**23** Vanadium (V) 50,942	**24** Chrome (Cr) 51,996	**25** Manganèse (Mn) 54,938	**26** Fer (Fe) 55,847	**27** Cobalt (Co) 58,933
37 Rubidium (Rb) 85,47	**38** Strontium (Sr) 87,62	**39** Yttrium (Y) 88,905	**40** Zirconium (Zr) 91,22	**41** Niobium (Nb) 92,906	**42** Molybdène (Mo) 95,94	**43°** Technétium (Tc) 98,91	**44** Ruthénium (Ru) 101,07	**45** Rhodium (Rh) 102,905
55 Césium (Cs) 132,905	**56** Baryum (Ba) 137,34	**57** Lanthane (La) 138,91	**58** Cérium (Ce) 140,12	**59** Praséodyme (Pr) 140,907	**60** Néodyme (Nd) 144,24	**61°** Prométhium (Pm) 145	**62** Samarium (Sm) 150,35	**63** Eucopium (Eu) 151,96
			72 Hafnium (Hf) 178,49	**73** Tantale (Ta) 180,948	**74** Tungstène (W) 183,85	**75** Rhénium (Re) 186,2	**76** Osmium (Os) 190,2	**77** Iridium (Ir) 192,2
87° Francium (Fr) 223	**88°** Radium (Ra) 226,05	**89°** Actinium (Ac) 227	**90°** Thorium (Th) 232,038	**91°** Protoactinium (Pa) 231	**92°** Uranium (U) 238,03	**93°** Neptunium (Np) 237	**94°** Plutonium (Pu) 242	**95°** Americium (Am) 243
			104° Rutherfordium (Rf) 259	**105°** Hahnium (Ha) 260				

								2 Hélium (He) 4,003
			5 Bore (B) 10,811	**6** Carbone (C) 12,011	**7** Azote (N) 14,007	**8** Oxygène (O) 15,999	**9** Fluor (F) 18,998	**10** Néon (Ne) 20,183
			13 Aluminium (Al) 26,982	**14** Silicium (Si) 28,086	**15** Phosphore (P) 30,974	**16** Soufre (S) 32,064	**17** Chlore (Cl) 35,453	**18** Argon (A) 39,948
28 Nickel (N) 58,71	**29** Cuivre (Cu) 63,54	**30** Zinc (Zn) 65,37	**31** Gallium (Ga) 69,72	**32** Germanium (Ge) 72,59	**33** Arsenic (As) 74,922	**34** Sélénium (Se) 78,96	**35** Brome (Br) 79,909	**36** Krypton (Kr) 83,80
46 Palladium (Pd) 106,4	**47** Argent (Ag) 107,870	**48** Cadmium (Cd) 112,40	**49** Indium (In) 114,82	**50** Étain (Sn) 118,69	**51** Antimoine (Sb) 121,75	**52** Tellure (Te) 127,60	**53** Iode (I) 126,904	**54** Xénon (Xe) 131,30
64 Gadolinium (Gd) 157,25	**65** Terbium (Tb) 158,924	**66** Dysprosium (Dy) 162,50	**67** Holmium (Ho) 164,930	**68** Erbium (Er) 167,26	**69** Thulium (Tm) 168,934	**70** Ytterbium (Yb) 173,04	**71** Lutécium (Lu) 174,97	
78 Platine (Pt) 195,09	**79** Or (Au) 196,967	**80** Mercure (Hg) 200,59	**81** Thallium (Tl) 204,37	**82** Plomb (Pb) 207,19	**83** Bismuth (Bi) 208,98	**84°** Polonium (Po) 210	**85°** Astate (At) 210	**86°** Radon (Rn) 222
96° Curium (Cm) 244	**97°** Berkélium (Bk) 245	**98°** Californium (Cl) 246	**99°** Einsteinium (Es) 253	**100°** Fermium (Fm) 255	**101°** Mendélévium (Md) 256	**102°** Nobélium (No) 255	**103°** Lawrencium (Lw) 257	

Tableau 6.1. Le tableau périodique des éléments. Les parties grisées constituent les deux séries de terres rares : les lanthanides et les actinides, ainsi désignées par leur premier élément respectif. Le poids atomique de chaque élément est inscrit en bas à droite de sa case. L'astérisque dénote un élément radioactif. Le numéro atomique de l'élément figure en haut de la case.

sous des angles caractéristiques et avec une symétrie tout aussi caractéristique. Cette régularité bien visible résulte d'un arrangement ordonné des atomes qui entrent dans sa structure. On avait des raisons de penser que la distance entre une couche d'atomes et sa voisine était de l'ordre de la longueur d'onde des rayons X. S'il en était bien ainsi, alors les cristaux devaient diffracter les rayons X.

Laue fit l'expérience et trouva que des rayons X passant à travers un cristal étaient effectivement diffractés pour former sur une plaque photographique une figure montrant qu'ils avaient les propriétés des ondes. La même année, le physicien anglais William Lawrence Bragg et son tout aussi illustre père, William Henry Bragg, mirent au point une méthode pour calculer avec précision la longueur d'onde d'un rayonnement X particulier à partir de sa figure de diffraction. Inversement, on a fini par utiliser les figures de diffraction des rayons X pour déterminer avec exactitude l'orientation des couches d'atomes diffractrices. Les rayons X ouvraient ainsi la porte à une nouvelle compréhension de la structure atomique des cristaux. Laue, en 1914, et les Bragg, en 1915, reçurent le prix Nobel de physique pour leurs travaux sur les rayons X.

Puis, en 1914, le jeune physicien anglais Henry Gwyn-Jeffreys Moseley détermina les longueurs d'onde des rayons X caractéristiques émis par divers métaux, et fit la découverte importante de la décroissance régulière de ces longueurs d'onde à mesure que le numéro atomique s'élevait.

Cette propriété affectait aux éléments des positions bien définies dans le tableau. Si deux éléments, supposés voisins dans le tableau, donnaient des rayons X dont les longueurs d'onde différaient du double de l'écart attendu, alors il devait y avoir une place entre les deux qui appartenait à un élément inconnu.

Il devenait maintenant possible d'attribuer aux éléments des nombres définis. Il y avait jusqu'alors toujours eu la possibilité d'une nouvelle découverte qui vienne s'intercaler dans la séquence et condamner au rebut tout système de numérotation adopté auparavant. Désormais il ne pouvait plus y avoir de case ignorée.

Les chimistes numérotèrent les éléments de 1, pour l'hydrogène, à 92, pour l'uranium. On s'aperçut que ces *numéros atomiques* jouaient un rôle en rapport avec la structure interne des atomes (voir le chapitre 7) et qu'ils étaient plus fondamentaux que les poids atomiques. Par exemple, les rayons X ont confirmé la validité de la décision de Mendeleïev lorsqu'il avait placé le tellurium (numéro atomique 52) avant l'iode (53), en dépit de son poids atomique plus élevé.

L'avantage du nouveau système de Moseley se révéla presque immédiatement. Le chimiste français Georges Urbain, après avoir découvert le *lutécium* (d'après l'ancien nom latin de la ville de Paris), annonça plus tard la découverte d'un autre élément qu'il appela *celtium.* Selon le système de Moseley, le lutécium était l'élément 71 et le celtium devait être le 72. Mais l'analyse des rayons X caractéristiques du celtium par Moseley révéla que le celtium n'était autre que du lutécium. Le véritable élément 72 ne fut pas découvert avant 1923, année où le physicien danois Dirk Coster et le chimiste hongrois Georg von Hevesy le détectèrent en laboratoire à Copenhague et le nommèrent *hafnium* (d'après le nom latinisé de Copenhague).

Moseley n'assista pas à cette confirmation de sa méthode ; il avait été tué à Gallipoli (Turquie) en 1915, à l'âge de vingt-huit ans – ce fut

certainement une des vies les plus précieuses parmi celles qui furent perdues au cours de la Première Guerre mondiale. Cette fin prématurée a probablement privé Moseley d'un prix Nobel. Le physicien suédois Karl Manne George Siegbahn continua les travaux de Moseley, découvrant de nouvelles séries de rayons X et déterminant avec précision les spectres X des divers éléments. Il a été récompensé par le prix Nobel de physique en 1924.

En 1925, Walter Noddack, Ida Tacke et Otto Berg, en Allemagne, ont comblé un autre trou dans le tableau périodique. Après trois ans d'études sur des minerais contenant les éléments de la même famille que celui qu'ils recherchaient, ils isolèrent l'élément 75 qu'ils appelèrent *rhénium,* pour évoquer le Rhin. Ne restaient que quatre trous : les éléments 43, 61, 85 et 87.

Deux décennies allaient être nécessaires pour débusquer ceux-ci. Les chimistes ne pouvaient le savoir à cette époque, mais ils avaient trouvé les derniers éléments stables. Les manquants étaient des éléments instables, si rares à présent sur Terre qu'à l'exception d'un d'entre eux, il faudrait les créer en laboratoire pour parvenir à les identifier. Et ceci est encore une belle histoire.

Les éléments radioactifs

L'IDENTIFICATION DES ÉLÉMENTS

La découverte des rayons X en 1895 poussa de nombreux chercheurs à étudier ce nouveau rayonnement terriblement pénétrant. Parmi ceux-là il y avait le physicien français Antoine Henri Becquerel. Son père, Alexandre Edmond (le physicien qui avait réalisé la première photographie du spectre solaire), s'était particulièrement intéressé à la *fluorescence,* ce rayonnement visible émis par certaines substances après avoir été exposées au rayonnement ultraviolet du Soleil.

Becquerel père avait en particulier étudié une substance fluorescente appelée *sulfate de potassium et d'uranium* (un composé dont la molécule contient un atome d'uranium). Henri se demandait si, dans le rayonnement de fluorescence du sulfate de potassium et d'uranium, figuraient des rayons X. Pour régler cette question, il fallait exposer le sulfate à la lumière solaire (dont la composante ultraviolette exciterait la fluorescence), le composé étant répandu sur une plaque photographique préalablement enveloppée dans un papier noir. Comme la lumière du Soleil ne pouvait pénétrer le papier noir, la plaque ne pouvait être impressionnée par celle-ci, mais si la fluorescence qu'elle excitait contenait des rayons X, alors ceux-ci *devaient* pénétrer le papier et voiler la plaque. Becquerel tenta et réussit cette expérience en 1896. Apparemment il y avait un rayonnement X dans la fluorescence. Becquerel parvint même à faire traverser des feuilles fines d'aluminium et de cuivre par ces prétendus rayons X et la question semblait donc réglée, car on ne connaissait aucun autre rayonnement capable d'en faire autant.

C'est alors que par chance, bien que ce ne fusse certainement pas l'avis de Becquerel à ce moment-là, survint une période de temps couvert. En

attendant le retour du soleil, Becquerel rangea ses plaques photos, saupoudrées de sulfate, dans un tiroir. Son impatience s'étant accrue après plusieurs jours d'attente, il décida de développer ses plaques de toute façon, dans l'idée que même sans éclairage solaire direct, quelques rayons X auraient pu être produits. Quand il vit les plaques développées, Becquerel vécut un de ces moments d'étonnement et de plaisir profonds dont tous les scientifiques rêvent. La plaque photographique était complètement voilée par un rayonnement intense ! Ceci était causé par quelque chose d'autre que la fluorescence ou la lumière solaire. Becquerel décida (ce que l'expérience confirma rapidement) que ce « quelque chose » était l'uranium dans le sulfate de potassium et d'uranium.

Les scientifiques, déjà agités par la découverte récente des rayons X, furent encore plus excités. Parmi les chercheurs qui se lancèrent immédiatement dans l'étude de cet étrange rayonnement de l'uranium, il y avait une jeune chimiste née en Pologne, Marie Sklodowska, qui avait épousé, un an auparavant, Pierre Curie, auteur de la découverte de la température de Curie (voir le chapitre 5).

En collaboration avec son frère Jacques, Pierre Curie avait découvert que certains cristaux, soumis à une pression, se chargeaient positivement sur une face et négativement sur la face opposée. Ce phénomène s'appelle la *piézo-électricité* (d'un mot grec signifiant « presser »). Marie Curie décida d'utiliser la piézo-électricité pour mesurer le rayonnement émis par l'uranium. Dans son appareil, le rayonnement ionisait l'air entre deux électrodes, et un faible courant s'écoulait, dont on estimait l'intensité en mesurant la pression qu'il fallait exercer sur un cristal pour produire un contre-courant de compensation. Cette méthode marchait si bien que Pierre Curie abandonna immédiatement ses propres travaux et consacra le reste de sa vie à seconder efficacement Marie.

C'est Marie Curie qui proposa le terme de radioactivité pour qualifier la propriété d'émission manifestée par l'uranium, et qui mit ce phénomène en évidence dans une deuxième substance radioactive, le thorium. D'importantes découvertes, faites aussi par d'autres chercheurs, se succédèrent rapidement. Les rayonnements d'autres substances radioactives se sont avérés encore plus pénétrants et énergétiques que les rayons X ; on les appelle maintenant les *rayons gamma*. On a découvert que des éléments radioactifs émettaient encore d'autres types de rayonnements, ce qui a conduit à des découvertes sur la structure interne de l'atome, mais le sujet sera traité dans un autre chapitre (voir le chapitre 7). C'est la découverte que des éléments radioactifs, en émettant leur rayonnement, se changeaient en d'autres éléments – une version moderne de la *transmutation* – qui est la plus importante pour l'objet de notre étude, les éléments.

Marie Curie s'est trouvée confrontée la première aux implications de ce phénomène, et ceci de façon accidentelle. En mesurant la teneur en uranium de la pechblende, pour déterminer si la quantité d'uranium dans des échantillons du minerai valait la peine de procéder au raffinage, elle et son mari eurent la surprise de trouver dans certains morceaux une radioactivité supérieure même à celle de l'uranium pur. Cela signifiait évidemment qu'il devait y avoir d'autres éléments radioactifs dans la pechblende. Ces éléments inconnus ne pouvaient figurer qu'en petite quantité, car l'analyse chimique ordinaire ne les détectait pas ; ils devaient donc être très radioactifs.

Très excités, les Curie se procurèrent des tonnes de pechblende, installèrent un atelier dans un appentis et – dans des conditions précaires et avec pour seul moteur leur indomptable enthousiasme – ils traquèrent des quantités infimes des nouveaux éléments dans cette masse de minerai noir. En juillet 1898 ils avaient isolé un soupçon d'une poudre noire quatre cents fois plus radioactive que la même quantité d'uranium.

Cette poudre contenait un nouvel élément que ses propriétés chimiques apparentaient au tellurium, au-dessous duquel il devait donc se trouver dans le tableau périodique des éléments. (On lui attribua plus tard le numéro atomique 84.) Les Curie le baptisèrent *polonium* en souvenir du pays natal de Marie.

Mais le polonium ne rendait compte que d'une partie de la radioactivité. D'autres travaux suivirent, et en décembre 1898 les Curie se trouvèrent en possession d'une préparation encore plus radioactive que le polonium. Elle contenait encore un autre élément, dont les propriétés étaient similaires à celles du baryum (et qui a finalement été placé sous le baryum et s'est révélé avoir le numéro atomique 88). En raison de son intense radioactivité, les Curie l'appelèrent *radium*.

Ils travaillèrent quatre ans de plus pour isoler une quantité visible de radium pur. Marie Curie présenta alors un résumé de ses travaux en guise de thèse de doctorat en 1903. C'est sans doute la thèse de doctorat la plus importante de toute l'histoire des sciences. Elle lui valut non seulement un, mais deux prix Nobel. Marie et son compagnon, conjointement avec Becquerel, se virent décerner le prix Nobel de physique en 1903 pour leur étude de la radioactivité ; et en 1911 Marie seule (son époux était mort écrasé par un fiacre en 1906) reçut le prix Nobel de chimie pour la découverte du polonium et du radium.

Le polonium et le radium sont bien plus instables que l'uranium et le thorium, autrement dit ils sont beaucoup plus radioactifs. Leurs atomes sont plus nombreux à se briser, toutes les secondes. Leur durée de vie est si brève que pratiquement tout le polonium et le radium de l'Univers auraient dû disparaître en un million d'années. Pourquoi en trouvons-nous encore sur cette Terre vieille de milliards d'années ? C'est parce que le radium et le polonium sont continuellement créés au cours des chaînes de désintégration de l'uranium et du thorium au plomb. Où il y a de l'uranium et du thorium, on a toutes les chances de trouver des traces de polonium et de radium. Ce sont des produits intermédiaires dans la transformation dont le produit final est le plomb.

Trois autres éléments instables, appartenant aux filiations de l'uranium et du thorium vers le plomb, ont été découverts par l'analyse minutieuse de la pechblende ou par des recherches sur les substances radioactives. En 1899, André Louis Debierne, conseillé par les Curie, chercha d'autres éléments dans la pechblende et trouva *l'actinium* (du mot grec qui signifie « rayon »), qui s'est révélé finalement avoir le numéro atomique 89. L'année suivante, le physicien allemand Friedrich Ernst Dorn montra que l'uranium, lorsqu'il se désintègre, produit un élément gazeux. Un gaz radioactif était une nouveauté ! Cet élément fut nommé *radon* (de radium et d'argon, ses cousins chimiques) et se vit attribuer le numéro atomique 86. Enfin, en 1917, deux groupes différents – Otto Hahn et Lise Meitner en Allemagne, Frederick Soddy et John Arnold Cranston en Angleterre – isolèrent de la pechblende l'élément 91 qu'ils nommèrent *protoactinium*.

À LA RECHERCHE DES ÉLÉMENTS ABSENTS

On était donc parvenu en 1925 au score de 88 éléments identifiés – 81 stables et 7 instables. La recherche des quatre absents – les numéros 43, 61, 85, 87 – devint fébrile.

Comme tous les éléments connus du numéro 84 au 92 étaient radioactifs, on s'attendait à ce que les 85 et 87 le soient eux aussi. D'autre part, les 43 et 61 étaient entourés d'éléments stables et il n'y avait aucune raison qu'il puisse ne pas en être de même pour eux ; on devait donc les touver dans la nature.

On supposait que l'élément 43 aurait des propriétés semblables au rhénium, au-dessus duquel il était dans le tableau, et qu'il se trouverait dans les mêmes minerais. L'équipe de Noddack, Tacke et Berg, qui avait découvert le rhénium, était d'ailleurs convaincue d'avoir aussi détecté des rayons X dont la longueur d'onde correspondait à l'élément 43. Ils annoncèrent donc sa découverte et la nommèrent *masurium,* d'après une région de Prusse-Orientale. Mais cette identification ne fut pas confirmée, et en matière scientifique une découverte n'en est effectivement une qu'après avoir été confirmée par au moins un groupe de chercheurs indépendant.

En 1926, deux chimistes de l'université d'Illinois annoncèrent la découverte de l'élément 61 dans des minerais contenant des éléments voisins (60 et 62) et le nommèrent *illinium.* La même année, deux chimistes italiens de l'université de Florence crurent avoir isolé le même élément qu'ils baptisèrent *florentium.* Mais les autres chimistes ne purent confirmer les travaux de ces groupes.

Quelques années plus tard, un physicien de l'institut polytechnique de l'Alabama annonça la découverte de traces des éléments 85 et 87 au moyen d'une nouvelle méthode d'analyse dont il était l'auteur ; il désigna ces éléments sous les noms de *virginium* et d'*alabamine,* respectivement d'après ses États natal et d'adoption. Mais ces découvertes ne furent pas plus confirmées.

La suite des événements devait bien montrer que les « découvertes » des éléments 43, 61, 85 et 87 étaient erronées.

C'est l'élément 43 qui fut d'abord identifié de façon certaine. Le physicien américain Ernest Orlando Lawrence, qui devait recevoir le prix Nobel de physique pour l'invention du cyclotron (voir le chapitre 7), fabriqua cet élément au moyen de son accélérateur en bombardant du *molybdène* (l'élément numéro 42) avec des particules à grande vitesse. Le matériau bombardé devenait radioactif, et Lawrence en envoya pour analyse au chimiste italien Emilion Gino Segrè, qui s'intéressait au problème de l'élément 43. Segrè et son collègue Carlo Perrier, après avoir séparé l'élément radioactif du molybdène, constatèrent que ses propriétés l'apparentaient au rhénium, mais que ce n'était pas du rhénium. Ils décidèrent que ce ne pouvait être que l'élément 43, et que celui-ci, contrairement à ses voisins du tableau, était radioactif. Comme il n'est pas produit par filiation d'un élément plus élevé, il n'en reste pratiquement plus dans la croûte terrestre, donc Noddack et ses collaborateurs faisaient certainement fausse route en croyant l'avoir trouvé. Segrè et Perrier eurent le privilège de nommer cet élément le *technétium,* d'après un mot grec signifiant « artificiel », car c'était le premier élément créé en laboratoire. En 1960 on avait accumulé assez de technétium pour en déterminer le

point de fusion – proche de 2 200° C. (Segrè devait recevoir plus tard le prix Nobel pour une tout autre découverte concernant encore quelque chose créé en laboratoire – voir le chapitre 7.)

L'élément 87 fut finalement découvert dans la nature en 1939. La chimiste française Marguerite Perey l'isola parmi les produits de filiation de l'uranium. L'élément 87 ne se présentait qu'en très petite quantité et ce sont les progrès techniques qui permirent de le trouver là où on l'avait vainement cherché. Marguerite Perey nomma cet élément le *francium.*

Comme le technétium, l'élément 85 fut produit à l'aide du cyclotron, par bombardement du *bismuth* (l'élément 83). En 1940, Segrè, Dale Raymond Corson et Kenneth Ross Mackenzie isolèrent l'élément 85 dans le laboratoire de l'université de Californie ; Segrè avait alors émigré d'Italie aux États-Unis. La Deuxième Guerre mondiale interrompit leurs travaux sur cet élément, mais ils y revinrent après la guerre et proposèrent en 1947 de le nommer *astate,* d'après un mot grec signifiant « instable ». On avait alors fini par trouver, comme pour le francium, de petites traces d'astate dans la nature, parmi les produits de filiation de l'uranium.

Entre-temps, le quatrième et dernier élément, le numéro 61, avait été identifié parmi les produits de fission de l'uranium, processus expliqué dans le chapitre 10. (Le technétium est également présent dans ces produits.) Trois chimistes du laboratoire d'Oak Ridge – J. A. Marinsky, L. E. Glendenin et Charles Dubois Coryell – ont isolé l'élément 61 en 1945. Ils l'ont nommé *prométhéum* en souvenir du demi-dieu Prométhée, qui avait volé le feu au Soleil pour le bien de l'humanité. Après tout, l'élément 61 avait été extrait du feu quasi solaire de la fournaise nucléaire.

Ainsi la liste des éléments, de 1 à 92, était enfin complète. Et pourtant, d'une certaine façon, l'étape la plus étrange de l'aventure ne faisait que commencer. En effet, les chercheurs avaient fait éclater la frontière du tableau périodique ; l'uranium n'en était pas la borne.

LES ÉLÉMENTS TRANSURANIENS

On avait effectivement commencé à chercher des éléments au-delà de l'uranium – les *éléments transuraniens* – dès 1934. Enrico Fermi avait trouvé qu'en bombardant un élément avec une particule subatomique qui venait d'être découverte, le *neutron* (voir chapitre 7), cet élément se voyait souvent transformé en son voisin de numéro atomique immédiatement supérieur. Pouvait-on, sur de l'uranium, bâtir un élément 93 – un élément totalement synthétique qui, pour autant qu'on le savait à l'époque, n'existait pas dans la nature ? Le groupe de Fermi s'attaqua à l'uranium avec des neutrons et il obtint un produit dont il pensa qu'il s'agissait de l'élément 93. Il l'appela *uranium X.*

Fermi reçut le prix Nobel de physique en 1938 pour ses travaux sur le bombardement neutronique. La vraie nature de sa découverte et ses conséquences pour l'humanité étaient, à cette époque, insoupçonnables. Comme cet autre Italien, Christophe Colomb, il avait trouvé non pas ce qu'il cherchait mais, sans s'en douter, quelque chose de beaucoup plus important.

Il suffit de dire ici qu'après avoir suivi bien des fausses pistes, on s'aperçut que Fermi n'avait pas créé un nouvel élément mais fendu l'atome d'uranium en deux parties presque égales. Lorsque les physiciens se mirent

à étudier ce processus en 1940, l'élément 93 surgit tout naturellement de leurs expériences. Dans le mélange d'éléments vestiges du bombardement de l'uranium par les neutrons, en subsistait un qui commença par défier les tentatives d'identification. Edwin McMillan, de l'université de Californie, eut alors l'idée que les neutrons libérés par la fission avaient peut-être converti des atomes d'uranium en un élément plus élevé, comme Fermi l'avait espéré. McMillan et Philip Abelson, un physico-chimiste, purent prouver que l'élément non identifié était en fait le numéro 93. Comme pour toutes les découvertes ultérieures, c'était la nature de sa radioactivité qui fournissait la preuve de son existence.

McMillan soupçonnait la présence d'un autre élément transuranien mélangé à l'élément 93. Le chimiste Glenn Theodore Seaborg et ses collaborateurs Arthur Charles Wahl et Joseph William Kennedy montrèrent bientôt que c'était le cas et que cet élément portait le numéro 94.

Comme l'uranium, limite supposée du tableau périodique, avait été nommé d'après la planète Uranus qui venait d'être découverte, les éléments 93 et 94 ont été baptisés *neptunium* et *plutonium* d'après les noms attribués aux planètes Neptune et Pluton découvertes entre-temps. Ces éléments ont aussi manifesté leur existence dans la nature ; on a trouvé plus tard des traces infimes de neptunium et de plutonium dans les minerais d'uranium. Ainsi l'uranium n'était finalement pas l'élément naturel le plus lourd.

Seaborg et un groupe de l'université de Californie, animé par Albert Ghiorso, se mirent à fabriquer d'autres éléments transuraniens, de plus en plus nombreux. En 1944, par bombardement du plutonium avec des particules subatomiques, ils ont créé les éléments 95 et 96, appelés respectivement *américium* (d'après Amérique) et *curium* (en hommage au couple Curie). Après avoir fabriqué des quantités utilisables d'américium et de curium, ils ont bombardé ces éléments et réussi à produire le numéro 97 en 1949 et le numéro 98 en 1950, qu'ils ont nommés le *berkélium* et le *californium*, d'après Berkeley et la Californie. Seaborg et McMillan ont partagé le prix Nobel de Chimie 1951 pour cette série de réussites.

Les éléments suivants ont été découverts dans des circonstances plus fracassantes. Les éléments 99 et 100 sont apparus dans l'explosion de la première bombe à hydrogène dans le Pacifique en novembre 1952. Bien que leur existence ait été détectée dans les débris de l'explosion, ces éléments n'ont été confirmés et baptisés qu'après leur fabrication, en petite quantité, par un groupe de l'université de Californie en 1955. On leur a attribué les noms d'*einsteinium* et de *fermium* en l'honneur d'Albert Einstein et Enrico Fermi, tous deux décédés quelques mois auparavant. Ce groupe a ensuite bombardé une petite quantité d'einsteinium pour former l'élément 101, qu'ils ont appelé *mendélévium* en hommage à Mendeleïev.

L'étape suivante a été le fruit d'une collaboration entre la Californie et l'institut Nobel en Suède. L'institut produisit une petite quantité d'élément 102 par un bombardement particulièrement compliqué. Pour honorer l'institut, l'élément a été nommé *nobélium*. Cet élément ayant été créé par des méthodes différentes de celles décrites par la première équipe de chercheurs, il s'est écoulé un certain délai avant que le nom de nobélium soit officiellement accepté.

En 1961 on a détecté quelques atomes d'élément 103 à l'université de Californie, auxquels on a attribué le nom de *lawrencium* en l'honneur de E.O. Lawrence, qui venait de disparaître. En 1964, un groupe de chercheurs

soviétiques animé par Georgii Nikolaevitch Flerov a annoncé la formation de l'élément 104 et, en 1967, la formation de l'élément 105. Dans ces deux cas, la validité des méthodes utilisées pour former ces éléments n'a pu être confirmée et les équipes américaines dirigées par Albert Ghiorso ont créé ces éléments par d'autres méthodes. L'équipe soviétique a baptisé le 104 *kurchatovium,* d'après Igor Vasilievitch Kurchatov, chef du groupe qui a mis au point la bombe atomique soviétique, mort en 1960. L'équipe américaine a nommé le 104 *rutherfordium* et le 105 *hahnium* en souvenir d'Ernest Rutherford et d'Otto Hahn, auteurs de découvertes fondamentales dans le domaine de la structure subatomique. On a annoncé des éléments jusqu'au numéro 109.

LES ÉLÉMENTS SUPERLOURDS

Chaque marche de cette ascension dans le domaine des transuraniens a été plus pénible à franchir que la précédente. A chaque étape, l'élément à mettre en évidence devenait plus difficile à accumuler et plus instable. Arrivé au mendélévium, il fallut se contenter de dix-sept atomes pour procéder à l'identification. Heureusement, en 1955, la technologie des détecteurs de rayonnements avait bénéficié de progrès considérables. Les chercheurs de Berkeley avaient connecté leurs détecteurs à une sonnette d'alarme ; ainsi chaque fois qu'un atome de mendélévium était formé, le rayonnement caractéristique que sa désintégration émettait ensuite proclamait l'événement par une sonnerie aussi triomphante qu'assourdissante. (Le service de sécurité incendie mit rapidement fin à cela.)

Les éléments plus élevés ont été détectés dans des conditions de pénurie encore plus grande. On peut détecter la présence d'un seul atome de l'élément cherché en dressant soigneusement la liste de ses produits de désintégration.

Quel intérêt y a-t-il à toujours tenter d'aller plus loin, hormis la joie de battre un record et de voir son nom immortalisé dans le grand registre des records pour avoir découvert un élément ? (Lavoisier, le plus grand des chimistes, n'est jamais parvenu à faire une telle découverte et cet échec le tracassait profondément.)

Il reste peut-être une découverte importante à faire. L'accroissement de l'instabilité lorsqu'on grimpe l'échelle des numéros atomiques n'est pas quelque chose d'uniforme. L'atome stable le plus complexe est le bismuth (83). Après celui-ci, les six éléments de 84 à 89 sont si instables que quelle qu'en soit la quantité qui aurait pu être présente au moment de la formation de la Terre, ils ont maintenant totalement disparu. Mais ils sont suivis, de façon surprenante, par le thorium (90) et l'uranium (92), qui sont presque stables. Du thorium et de l'uranium primordiaux qui se trouvaient sur la Terre à sa formation, il reste aujourd'hui respectivement 80 % et 50 %. Les physiciens ont élaboré des théories de la structure nucléaire pour expliquer ces irrégularités (comme je l'expliquerai dans le chapitre suivant), et si l'on en croit ces théories, les éléments 110 et 114 devraient être plus stables que leurs numéros atomiques ne le laissaient supposer. Obtenir ces éléments présente donc un intérêt considérable pour toutes ces théories *.

---

* Ajoutons, au risque de ternir la pureté de ces motivations, que les superlourds sont très recherchés par les militaires, car les théories prédisent un taux de production de neutrons beaucoup plus élevé lors de la fission de ces éléments, ce qui permettrait, en abaissant leur masse critique, d'en faire des mini-bombes thermonucléaires. *(N.D.T.)*

En 1976, certains ont prétendu que des *halos* (des traces circulaires sombres dans les micas) pourraient dénoter la présence de ces éléments *superlourds*. Les halos sont causés par le rayonnement de petites inclusions de thorium et d'uranium, mais il y a quelques halos géants qui doivent provenir d'atomes radioactifs plus énergétiques, néanmoins assez stables pour avoir subsisté jusqu'à maintenant. Ces déductions n'ont malheureusement pas été acceptées par le monde scientifique et l'idée a été abandonnée. Les chercheurs cherchent toujours.

Les électrons

Lorsque Mendeleïev et ses contemporains s'aperçurent qu'ils pouvaient disposer les éléments dans un tableau périodique composé de familles de substances dont les propriétés étaient analogues, ils n'avaient aucune explication à présenter pour ces regroupements d'éléments et ces similitudes de propriétés. Finalement une raison simple et convaincante s'est faite jour, mais seulement après une longue série de découvertes qui n'avaient, de prime abord, rien à voir avec la chimie.

Tout a commencé par l'étude de l'électricité. Faraday réalisait toutes les expériences qu'il pouvait imaginer avec l'électricité, et il tenta entre autres de créer une décharge électrique dans le vide. Il ne pouvait obtenir un vide suffisant pour cette expérience. Mais en 1854 un souffleur de verre allemand, Heinrich Geissler, inventa une pompe à vide adéquate et fabriqua un tube de verre dans lequel se trouvaient deux électrodes métalliques, dans un vide d'une qualité sans précédent. En produisant des décharges électriques dans le *tube de Geissler*, les expérimentateurs remarquaient une lueur verte sur la paroi du tube en face de l'électrode négative. Le physicien allemand Eugen Goldstein suggéra en 1876 que cette lueur verte était due à l'impact sur le verre d'une sorte de rayonnement émis par l'électrode négative, appelée *cathode* par Faraday. Goldstein baptisa ce rayonnement du nom de *rayons cathodiques*.

Les rayons cathodiques étaient-ils une forme de rayonnement électromagnétique ? C'est ce que pensait Goldstein, mais d'autres, comme le physicien anglais William Crookes, n'étaient pas de cet avis : il devait y avoir un flux de particules d'une espèce ou d'une autre. Crookes mit au point des versions améliorées du tube de Geissler (appelées *tubes de Crookes*), avec lesquelles il parvint à montrer que les rayons cathodiques étaient déviés par un aimant. Ils étaient ainsi probablement constitués de particules chargées électriquement, et le sens de la déviation indiquait qu'elles étaient négatives.

En 1897, le physicien Joseph John Thomson régla définitivement la question en montrant que les rayons cathodiques étaient aussi déviés par des charges électriques. Quelle était donc la nature de ces « particules » cathodiques ? Les seules particules chargées négatives connues à l'époque étaient les ions atomiques négatifs. Les expériences montraient que les particules des rayons cathodiques ne pouvaient pas être des ions ; en effet leur déviation dans un champ magnétique était si importante qu'elles devaient avoir une charge électrique incroyablement élevée ou alors être très légères, avec une masse inférieure au millième de la masse d'un atome

d'hydrogène. C'est cette dernière hypothèse qui fut confirmée par l'expérience. Les physiciens avaient déjà eu l'idée que le courant électrique était constitué de particules, et les particules des rayons cathodiques furent considérées comme les ultimes particules d'électricité. On les appela *électrons* – un nom proposé en 1891 par le physicien irlandais George Johnstone Stoney. On détermina finalement que la masse de l'électron valait 1/1 836 fois la masse de l'atome d'hydrogène. (Thomson reçut le prix Nobel de physique en 1906 pour avoir établi l'existence de l'électron.)

La découverte de l'électron suggéra immédiatement qu'il pouvait s'agir là d'une sous-particule de l'atome – en d'autres termes, les atomes ne constituaient plus l'ultime unité de matière indivisible imaginée par Démocrite et John Dalton.

La pilule était difficile à avaler, mais l'ensemble des preuves convergeaient inexorablement. Thomson emporta la conviction en éclairant une plaque métallique par une lumière ultraviolette et en montrant que les particules qui en jaillissaient *(l'effet photo-électrique)* étaient identiques aux électrons des rayons cathodiques. Il fallait bien que ces photo-électrons aient été arrachés aux atomes du métal.

LA PÉRIODICITÉ DE LA CLASSIFICATION

Étant donné la facilité avec laquelle on pouvait extraire les électrons d'un atome (non seulement par l'effet photo-électrique, mais par bien d'autres procédés), il était tout naturel d'en conclure qu'ils devaient se trouver près de la surface de celui-ci. Il devait donc y avoir dans l'atome une région chargée positivement, pour compenser les charges négatives des électrons, car l'atome est globalement neutre. C'est alors que les chercheurs commencèrent à approcher de la solution du mystère du tableau périodique.

Extraire un électron d'un atome requiert un peu d'énergie. Inversement, lorsqu'un électron retombe dans l'emplacement ainsi libéré dans l'atome, il doit *céder* la même quantité d'énergie. (La nature est d'ordinaire symétrique, surtout lorsqu'il s'agit d'énergie.) Cette énergie est produite sous forme de rayonnement électromagnétique. Mais comme l'énergie de ce rayonnement est fonction de sa longueur d'onde, la longueur d'onde du rayonnement émis par un électron tombant dans un atome donné va indiquer la force avec laquelle l'électron est maintenu dans cet atome. L'énergie du rayonnement est une fonction décroissante de sa longueur d'onde : plus grande est l'énergie, plus courte est la longueur d'onde.

C'est ainsi que nous arrivons à la découverte de Moseley : les métaux (c'est-à-dire les éléments les plus lourds) émettent des rayons X avec des longueurs d'onde caractéristiques qui décroissent de façon régulière lorsqu'on progresse dans la classification périodique. Chaque élément, semble-t-il, retient ses électrons un peu plus fortement que le précédent – autrement dit, la charge positive intérieure à cet élément est plus grande.

En supposant que chaque unité de charge positive correspondait à la charge négative d'un électron, il s'ensuivait que l'atome d'un élément devait avoir un électron de plus que son prédecesseur dans la classification. La façon la plus simple de s'expliquer cette classification consistait alors à supposer que le premier élément, l'hydrogène, avait une unité de charge positive et un électron, le second élément, l'hélium, deux charges positives

et deux électrons, le troisième, le lithium, trois charges positives et trois électrons, et ainsi de suite jusqu'à l'uranium avec 92 charges positives et 92 électrons. Il s'est ainsi avéré que les numéros atomiques des éléments représentaient exactement le nombre d'électrons que contiennent leurs atomes neutres.

Encore un indice majeur, et les physiciens atomistes tenaient la réponse à l'énigme de la périodicité du tableau périodique. On s'aperçut que le rayonnement électrique d'un élément donné n'était pas nécessairement restreint à une seule longueur d'onde ; l'atome pouvait émettre son rayonnement sur deux, trois, quatre, et même plus, longueurs d'onde différentes. Ces rayonnements, organisés en groupe, ont été nommés *série K, série L, série M* et ainsi de suite. Les chercheurs sont arrivés à la conclusion que les électrons étaient disposés en *couches* autour du cœur de l'atome chargé positivement. Les électrons de la couche la plus intérieure sont les plus liés et leur extraction exige plus d'énergie. Un électron qui tombe dans cette couche va émettre le rayonnement le plus énergétique, ou de longueur d'onde la plus courte, c'est-à-dire dans la série K. Les électrons de la couche immédiatement supérieure émettent les rayonnements de la série L ; la couche suivante produit la série M, etc. Ainsi, on a appelé les couches électroniques *couche K, couche L, couche M,* etc.

En 1925, le physicien autrichien Wolfgang Pauli a proposé son *principe d'exclusion* pour rendre compte de la répartition des électrons dans chaque couche, deux électrons ne pouvant d'après ce principe avoir exactement les mêmes valeurs pour leurs *nombres quantiques.* Ce succès lui a valu le prix Nobel de physique en 1945.

LES GAZ RARES

C'est en 1916 que le chimiste américain Gilbert Newton Lewis a fondé le comportement chimique et la similitude des propriétés de quelques-uns des éléments les plus simples sur la base de leur structure en couches. Pour commencer, la limitation de l'occupation de la couche la plus intérieure à deux électrons (par ce qui devait devenir le principe d'exclusion de Pauli) était amplement prouvée. L'hydrogène n'a qu'un électron, la couche n'est donc pas pleine. La tendance de l'atome est de compléter cette couche K, ce qu'il peut réaliser de plusieurs façons. Par exemple, deux atomes d'hydrogène peuvent mettre en commun leurs uniques électrons et, en partageant ces deux électrons, combler mutuellement leur couche K. Ainsi l'hydrogène gazeux se rencontre-t-il presque toujours sous forme de paires d'atomes – les molécules d'hydrogène. Séparer et libérer les deux atomes pour faire de l'hydrogène atomique requiert pas mal d'énergie. Irving Langmuir, de la compagnie General Electric, qui travaillait indépendamment sur une théorie du même type reliant électrons et comportement chimique, a mis au point une application pratique de cette forte tendance de l'hydrogène à maintenir le remplissage de sa couche électronique. Langmuir a réalisé un *chalumeau à hydrogène atomique* en faisant passer un jet d'hydrogène à travers un arc électrique pour briser les molécules et libérer les atomes ; en se recombinant après la traversée de l'arc, les atomes libèrent l'énergie que leur séparation avait absorbée, ce qui permet d'atteindre des températures allant jusqu'à 3 400° C !

Dans l'hélium, l'élément numéro 2, les deux électrons remplissent la couche K ; les atomes d'hélium sont donc stables et ne se combinent pas à d'autres atomes. Avec le lithium, l'élément numéro 3, nous trouvons deux électrons remplissant la couche K, tandis que le troisième inaugure la couche L. Les éléments successifs ajoutent des électrons un à un dans cette couche : le béryllium comporte deux électrons dans la couche L, le bore trois, le carbone quatre, l'azote cinq, l'oxygène six, le fluor sept, et le néon huit. La capacité limite de la couche L est de huit électrons, comme Pauli l'a montré, et le néon correspond donc à l'hélium en ceci que sa couche électronique externe est pleine. Effectivement, c'est aussi un gaz inerte dont les propriétés sont similaires à celles de l'hélium.

Tout atome dont la couche extérieure est incomplète tend à se combiner à d'autres atomes de façon à se retrouver avec une couche extérieure complète. Par exemple, l'atome de lithium cède volontiers son unique électron de la couche L pour que sa couche extérieure devienne la couche K pleine, tandis que le fluor cherche à se saisir d'un électron pour l'ajouter aux sept de sa couche L et compléter celle-ci. Ainsi le lithium et le fluor ont une affinité l'un pour l'autre ; lorsqu'ils se combinent, le lithium cède son électron L au fluor pour compléter la couche L de celui-ci. Comme les charges intérieures positives des atomes ne varient pas, le lithium avec un électron de moins a alors une charge résultante positive, tandis que le fluor, avec un électron en plus, a une charge négative. L'attraction mutuelle des charges opposées maintient ensemble les deux ions. Le composé s'appelle le fluorure de lithium (voir la figure 6.1).

Les électrons de la couche L sont aussi bien partageables que transmissibles. Par exemple, chacun des deux atomes de fluor peut partager l'un de ses électrons avec l'autre atome, de sorte que chaque atome voit sa couche L remplie de huit électrons, en comptant les deux électrons mis en commun. De même, deux atomes d'oxygène vont mettre en tout quatre électrons dans le pot commun pour compléter leur couche L ; et deux atomes d'azote vont se partager six électrons. Le fluor, l'oxygène et l'azote forment tous des molécules à deux atomes.

L'atome de carbone, avec seulement quatre électrons dans sa couche L, va partager chacun d'eux avec un atome d'hydrogène différent, remplissant ainsi les couches K de quatre atomes d'hydrogène et comblant à son tour sa propre couche L en partageant *leurs* électrons. Cet accommodement stable est celui de la molécule de méthane CH_4.

De la même façon, un atome d'azote peut partager des électrons avec trois atomes d'hydrogène pour constituer l'ammoniac, un atome d'oxygène partager des électrons avec deux atomes de carbone pour faire du dioxyde de carbone (ou oxyde de carbone), et ainsi de suite. Presque tous les composants formés à partir d'éléments figurant dans la première partie du tableau périodique peuvent s'expliquer par cette tendance au remplissage de la couche extérieure en acceptant, en cédant, ou même en partageant des électrons.

L'élément suivant le néon, le sodium, a onze électrons dont le dernier doit inaugurer une troisième couche. Viennent ensuite le magnésium, avec deux électrons dans la couche M, l'aluminium avec trois électrons, le silicium quatre, le phosphore cinq, le soufre six, le chlore sept et l'argon huit.

Chaque élément de ce groupe correspond à l'un des éléments de la série précédente. L'argon, avec huit électrons dans la couche M, est semblable

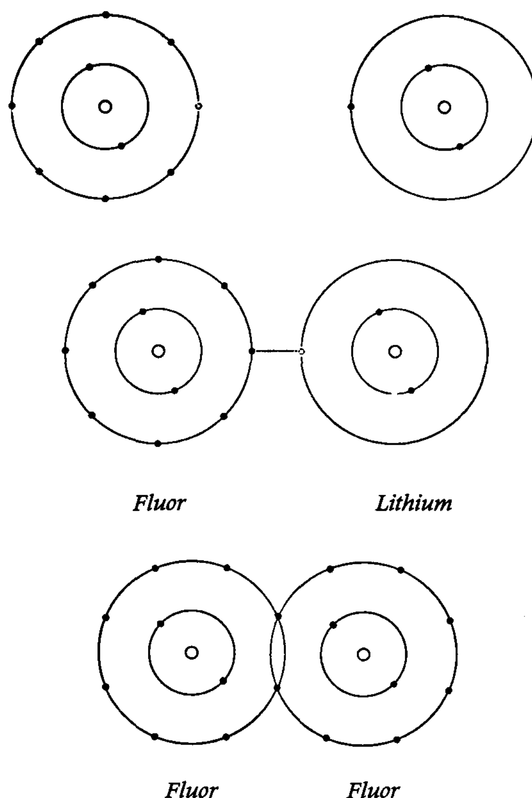


Figure 6.1. Transfert et partage des électrons. Le lithium transfère l'électron de sa couche externe au fluor pour constituer le fluorure de lithium ; chaque atome voit ainsi sa couche externe complétée. Dans la molécule de fluor, deux électrons sont mis en commun pour remplir les couches externes des deux atomes.

au néon (huit électrons dans la couche L) et c'est un gaz inerte. Le chlore, avec sept électrons dans sa couche extérieure, ressemble étroitement au fluor par ses propriétés chimiques. De même, le silicium ressemble au carbone, le sodium au lithium, et ainsi de suite.

Il en va de même dans tout le tableau périodique. Le comportement chimique de chaque élément dépendant de la configuration électronique de sa dernière couche, tous les éléments dont la couche externe comporte, disons un électron, réagiront chimiquement de façon tout à fait semblable. Ainsi, tous les éléments de la première colonne du tableau périodique – le lithium, le sodium, le potassium, le rubidium, le césium et même le francium radioactif – ont des propriétés chimiques remarquablement similaires. Le lithium comporte un électron dans la couche L, le sodium un électron dans la couche M, le potassium un dans la couche N, le rubidium un dans la couche O, le césium un dans la couche P, et le francium un dans la couche Q. Les éléments avec sept électrons dans la

couche externe – le fluor, le chlore, le brome, l'iode et l'astate – se ressemblent aussi. C'est la même chose pour la dernière colonne du tableau – le groupe à couches fermées composé de l'hélium, du néon, de l'argon, du krypton, du xénon et du radon.

L'idée de Lewis et Langmuir fonctionne si bien qu'elle permet encore, dans sa forme originale, de rendre compte des différences les plus fondamentales et évidentes dans le comportement des éléments. Mais il existe quand même des comportements pas aussi simples que l'on pouvait s'y attendre.

Par exemple, chacun des gaz rares – hélium, néon, argon, krypton, xénon et radon – a huit électrons dans sa couche externe (à l'exception de l'hélium dont la seule couche est comblée par deux électrons), et nous avons là la situation la plus stable possible. Pour les atomes de ces éléments, la tendance à céder ou prendre des électrons est minimale et il en va de même pour leur tendance à s'entremettre dans des réactions chimiques. Ces gaz ont tout pour être *inertes.*

Et pourtant, une « tendance au minimum » n'est pas tout à fait la même chose que « pas de tendance du tout », et la plupart des chimistes oublièrent cette vérité et firent comme s'il était strictement impossible pour les gaz inertes de former des composés. Mais ce n'était pas le cas de tous les chimistes. Dès 1932, le chimiste américain Linus Pauling, après avoir étudié l'énergie nécessaire à l'extraction des électrons des divers éléments, avait remarqué que tous les éléments sans exception, même les gaz inertes, pouvaient être dépouillés de leurs électrons. Cette soustraction nécessite toutefois plus d'énergie dans le cas des gaz inertes que pour leurs voisins dans le tableau périodique.

La quantité d'énergie requise pour extraire des électrons des éléments d'une famille donnée décroît en fonction du poids atomique, et les exigences des gaz inertes les plus lourds, le xénon et l'argon, ne sont pas particulièrement élevées. Il n'est pas plus difficile d'arracher un électron d'un atome de xénon, par exemple, que d'un atome d'hydrogène.

Pauling prédit ainsi que les gaz inertes les plus lourds pourraient former des composés chimiques avec des éléments particulièrement réceptifs aux électrons. Le fluor est l'élément le plus avide d'électrons et semblait donc le candidat naturel.

Mais le radon, le plus lourd des gaz inertes, est radioactif et n'existe qu'en très petite quantité, à l'état de traces. Le xénon, l'élément suivant par ordre de poids décroissant, est quant à lui stable et se trouve en petite quantité dans l'atmosphère. Il y avait donc intérêt à chercher à former un composé avec le xénon et le fluor. Mais rien ne fut fait pendant trente ans, à cause surtout du prix du xénon et des difficultés de manipulation du fluor, et les chimistes estimaient qu'il y avait mieux à faire que de chasser ce feu follet.

Finalement en 1962, Neil Bartlett, chimiste canadien d'origine britannique, alors qu'il travaillait sur un nouveau composé, l'hexafluorure de platine (PtF_6), réalisa que ce dernier était particulièrement avide d'électrons, presque autant que le fluor lui-même. Ce composé pouvait prendre des électrons à l'oxygène, un élément préférant normalement gagner des électrons plutôt qu'en perdre. Si le PtF_6 était capable d'arracher des électrons à l'oxygène, il devait pouvoir en faire autant avec le xénon. Il ne restait plus qu'à tenter l'expérience, et l'existence du platino-fluorure de xénon ($XePtF_6$) fut proclamée.

Immédiatement, d'autres chimistes emboîtèrent le pas et un certain nombre de composés du xénon avec le fluor, avec l'oxygène ou avec les deux furent formés, le plus stable étant le difluore de xénon (XeF_2). On a également formé un composé du krypton et du fluor, le tétrafluorure de krypton (KrF_4), ainsi qu'un fluorure de radon. Les composés avec l'oxygène sont, par exemple, l'oxytétrafluorure de xénon ($XeOF_4$), l'acide xénique (H_2XeO_4) et le xénate de sodium (Na_4XeO_6). Le plus intéressant de tous est peut-être le trioxyde de xénon (XeO_3), qui explose avec une facilité dangereuse. Les gaz inertes plus légers – l'argon, le néon et l'hélium – sont moins enclins à partager leurs électrons que leurs cousins lourds, et restent encore inertes face à toutes les sollicitations des chimistes.

Les chimistes se sont vite remis du choc causé par la découverte que les gaz inertes pouvaient consentir à former des composés ; après tout, ces composés s'intègrent très bien dans le panorama. C'est pourquoi on éprouve maintenant une certaine réticence à qualifier ces gaz d'inertes. On préfère dire les *gaz nobles* *, et l'on parle de *composés des gaz nobles* et de *chimie des gaz nobles.* (Je trouve que c'est aller de mal en pis. Ces gaz sont après tout encore inertes, même s'ils ne le sont pas totalement. Le terme de *noble,* impliquant dans ce contexte « distant » ou « répugnant à se commettre avec le vulgaire », me semble aussi peu approprié qu'*inerte* et, de plus, ne convient guère dans une société démocratique.)

LES TERRES RARES

Outre le fait que l'on appliquait trop rigidement la théorie de Lewis et Langmuir aux gaz rares, celle-ci était pratiquement inutilisable pour la plupart des éléments de numéro atomique supérieur à 20. Elle nécessitait en particulier quelques améliorations pour pouvoir traiter un phénomène très étrange dans le tableau périodique ayant trait aux *terres rares,* les éléments 57 à 71 inclus.

Revenons un peu en arrière, aux premiers chimistes pour qui une substance insoluble dans l'eau et insensible à la chaleur était une *terre* (un vestige du point de vue grec selon lequel la terre était un élément). Parmi ces substances, on trouvait ce que nous appellerions aujourd'hui l'oxyde de calcium, l'oxyde de magnésium, le dioxyde de silicium, l'oxyde de fer, l'oxyde d'aluminium, etc. – des composés qui constituent en fait 90 % de l'écorce terrestre. Les oxydes de calcium et de magnésium sont légèrement solubles et présentent, en solution, des propriétés *alcalines* (c'est-à-dire opposées à celles des acides) et furent donc baptisés *terres alcalines ;* lorsque Humphry Davy isola le calcium et le magnésium de ces terres rares, ceux-ci furent appelés *métaux alcalino-terreux.* On a finalement accolé ce nom à tous les éléments figurant dans la même colonne du tableau périodique que le magnésium et le calcium, à savoir le béryllium, le strontium, le baryum et le radium.

L'énigme à laquelle j'ai fait allusion a commencé en 1794, lorsqu'un chimiste finlandais, Johan Gadolin, décida, après avoir étudié une roche

---

* ... chez les Anglo-Saxons. En français, on parle plutôt de *gaz rares,* expression qui a le mérite de la justesse, de la transparence et de la neutralité idéologique. Elle n'évoque par contre rien des caractéristiques physiques de ces gaz, leur inertie n'étant qu'une des causes, lointaine, de leur rareté sur notre planète. *(N.D.T.)*

bizarre trouvée près du village d'Ytterby, en Suède, qu'il s'agissait d'une nouvelle « terre ». Gadolin donna à cette terre rare le nom d'*yttria,* d'après Ytterby. Plus tard, le chimiste allemand Martin Heinrich Klaproth s'aperçut que l'yttria pouvait être séparée en deux terres, l'une pour laquelle il conserva le nom d'yttria, tandis qu'il nomma l'autre *céria* (d'après l'astéroïde Cérès qui venait d'être découvert). Mais le chimiste suédois Carl Gustav Mosander devait encore décomposer celles-ci en un ensemble de terres différentes. Ces terres s'avérèrent finalement être des oxydes de nouveaux éléments métalliques maintenant appelés *terres rares.* En 1907, quatorze de ces éléments avaient été identifiés. Par ordre de poids atomique croissant, ce sont :

- le lanthane (d'après un mot grec signifiant « caché »)
- le cérium (d'après Cérès)
- le praséodyme (du grec pour « jumeau vert », à cause d'une raie verte de son spectre)
- le néodyme (« nouveau jumeau »)
- le samarium (d'après « samarskite », le minerai dans lequel il fut trouvé)
- l'europium (d'après Europe)
- le gadolinium (en l'honneur de Johan Gadolin)
- le terbium (d'après Ytterby)
- le dysprosium (d'après un mot grec signifiant « d'accès difficile »)
- l'holmium (d'après Stockholm)
- l'erbium (d'après Ytterby)
- le thulium (d'après le nom ancien de Thulé, désignant la Scandinavie)
- l'ytterbium (d'après Ytterby)
- le lutécium (d'après Lutèce, l'ancien nom de Paris).

Les propriétés de leurs spectres de rayons X permirent d'assigner à ces éléments les numéros atomiques 57 (pour le lanthane) à 71 (pour le lutécium). Comme je l'ai déjà raconté, il y avait un trou à 61, jusqu'à ce que l'élément manquant, le prométhéum, émerge de la fission de l'uranium, ce qui porte le nombre de ces éléments à quinze.

Mais le problème avec ces terres rares est qu'il semble impossible de les ranger dans la classification périodique. Il est heureux que l'on en ait connu seulement quatre lorsque Mendeleïev proposa sa classification ; si elles avaient toutes été connues, le tableau aurait sans doute été beaucoup trop confus pour être accepté. L'ignorance est parfois une bénédiction, même en science.

Le premier des métaux rares, le lanthane, s'apparente tout à fait à l'yttrium, le numéro 39, l'élément situé au-dessus dans le tableau (figure 6.2). (L'yttrium, bien qu'on le trouve dans les mêmes minerais que les terres rares et qu'il ait des propriétés similaires, n'est pas un métal des terres rares. Mais il doit encore son nom à Ytterby. Quatre éléments honorent ce village, ce qui est un peu exagéré.) La confusion commence avec la terre rare qui suit le lanthane, à savoir le cérium, et qui devrait ressembler à l'élément suivant l'yttrium, c'est-à-dire le zirconium. Mais il n'en fait rien, et ressemble plutôt encore à l'yttrium. Et il en est de même pour les quinze éléments des terres rares : ils se ressemblent tous et s'apparentent fortement à l'yttrium (ils sont en fait chimiquement si semblables que dans les premiers temps, on ne pouvait les séparer que par des procédés particulièrement fastidieux), mais ne sont reliés à aucun autre élément précédent dans la table. Il faut sauter tout le groupe des terres rares et parvenir à l'hafnium, l'élément 72, pour trouver un parent du zirconium qui suit l'yttrium.

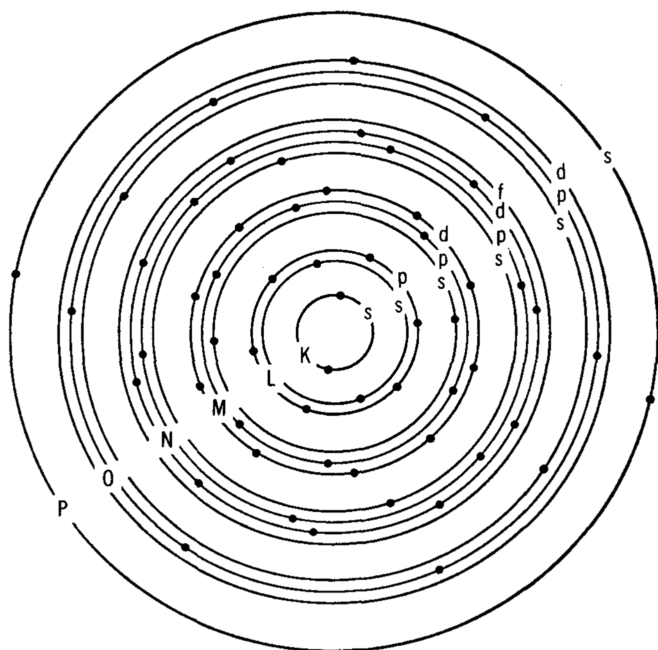


Figure 6.2. Les couches électroniques du lanthane. Remarquez le vide de la quatrième sous-couche de la couche N.

Bloqués par cette situation, les chimistes ne pouvaient faire mieux que regrouper tous les éléments des terres rares dans une case au-dessous de l'yttrium et les énumérer individuellement dans un additif à la table.

LES ÉLÉMENTS DE TRANSITION

La solution de l'énigme vint finalement de détails ajoutés au modèle de structure des éléments en couches d'électrons, de Lewis et Langmuir.

C.R. Bury suggéra en 1921 que la capacité des couches n'était pas limitée à huit électrons par couche. Huit suffisaient toujours pour remplir la couche externe. Mais une couche pouvait contenir plus d'électrons lorsqu'elle n'était pas à l'extérieur. Quand une couche se constituait au-dessus d'une autre, les couches internes pouvaient accueillir plus d'électrons, et chaque couche successive pouvait abriter plus d'électrons que la précédente. Ainsi la capacité totale de la couche L devait être de huit, la couche M de dix-huit, la couche N de trente-deux, et ainsi de suite en une série des doubles des carrés successifs (c'est-à-dire 2×1, 2×4, 2×9, 2×16, etc.).

Ce point de vue s'appuyait sur une étude détaillée des spectres des éléments. Le physicien danois Niels Henrik David Bohr avait montré que chaque couche électronique était constituée de sous-couches dont les niveaux d'énergie différaient légèrement. En montant dans les couches successives, les sous-couches étaient de plus en plus séparées, de sorte que les couches finissaient par s'imbriquer. La sous-couche la plus extérieure

d'une couche interne (disons la couche M) pouvait alors se trouver pour ainsi dire plus éloignée du centre que la sous-couche la plus intérieure de la couche suivante (disons la couche N). De ce fait, la sous-couche intérieure de la couche N pouvait s'emplir d'électrons tandis que la sous-couche externe de la couche M restait vide.

Un exemple va éclaircir cela. D'après la théorie, la couche M se divise en trois sous-couches dont les capacités sont respectivement de deux, six et dix électrons, soit dix-huit au total. Mais l'argon, avec huit électrons dans sa couche M, ne remplit que ses deux sous-couches internes. Et en fait, la troisième sous-couche M, la plus extérieure, ne va pas recevoir l'électron suivant dans le processus d'édification des éléments, car elle se trouve au-delà de la sous-couche la plus intérieure de la couche N : dans le potassium, l'élément qui suit l'argon, le dix-neuvième électron ne va pas dans la sous-couche M extérieure, mais dans la sous-couche N intérieure. Le potassium, avec un électron dans sa couche N, ressemble au sodium qui a un électron dans sa couche M. L'élément suivant, le calcium (20), a deux électrons dans la couche N et ressemble au magnésium avec ses deux électrons dans la couche M. Mais à présent la sous-couche la plus intérieure de la couche N, qui n'a de place que pour deux électrons, est pleine. Les prochains électrons ajoutés peuvent commencer à remplir la sous-couche extérieure de la couche M, qui était restée intacte. Le scandium (21) inaugure ce processus que le zinc (30) achève. Dans le zinc, la sous-couche extérieure de la couche M a finalement son plein de dix électrons. Les trente électrons du zinc se répartissent comme suit : deux dans la couche K, huit dans la couche L, dix-huit dans la couche M et deux dans la couche N. A ce stade, les électrons peuvent reprendre le remplissage de la couche N. L'électron suivant dote la couche N de trois électrons et forme le gallium (31), semblable à l'aluminium avec ses trois électrons dans la couche M.

Finalement, ces éléments 21 à 30, formés par le remplissage progressif d'une sous-couche temporairement négligée, sont des *éléments de transition.* Notons que le calcium ressemble au magnésium et le gallium à l'aluminium. Mais le magnésium et l'aluminium sont des voisins dans le tableau (ils portent les numéros 12 et 13), à la différence du calcium (20) et du gallium (31). Entre ceux-ci s'introduisent les éléments de transition qui viennent compliquer le tableau périodique.

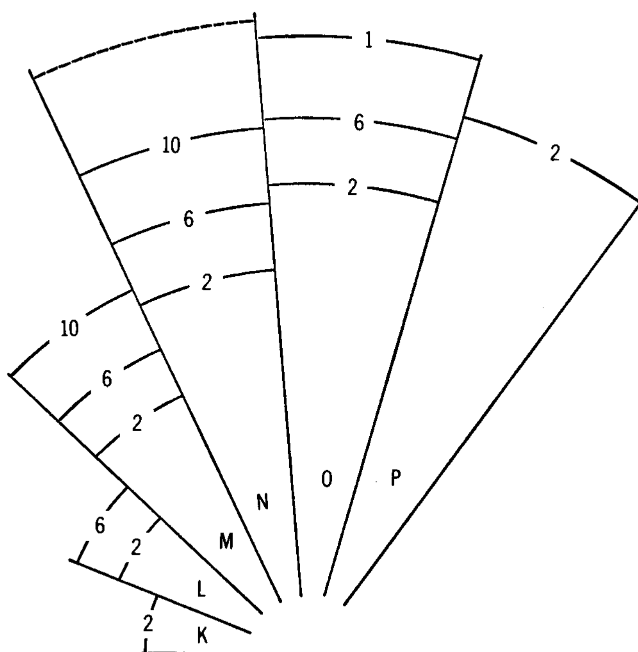
La couche N est plus spacieuse que la couche M ; elle se divise en quatre sous-couches, au lieu de trois, qui peuvent héberger respectivement deux, six, dix et quatorze électrons. L'élément 36, le krypton, comble les deux sous-couches les plus intérieures de la couche N, mais c'est alors que la sous-couche interne de la couche O intervient, et les électrons doivent remplir celle-ci avant de poursuivre avec les deux sous-couches extérieures de la couche N. Le rubidium (37), l'élément qui suit le krypton, reçoit son trente-septième électron dans la couche O. Le strontium (38) complète le remplissage de la sous-couche O à deux électrons. A partir de là, une nouvelle série d'éléments de transition se consacrent au remplissage de la troisième sous-couche N qui avait été négligée. Cela s'achève avec le cadmium (48) ; la quatrième sous-couche N, la plus externe, est alors omise, et les électrons viennent occuper la deuxième sous-couche O, jusqu'au xénon (54).

Mais la quatrième sous-couche N doit encore passer son tour ; l'empiétement des couches est devenu si important que même la couche P

vient intercaler une sous-couche à remplir avant la dernière de la couche N. Après le xénon, viennent le césium (55) et le baryum (56), avec un et deux électrons respectivement dans la couche P. Mais ce n'est pas encore le tour de la couche N : étonnamment, le cinquante-septième électron va dans la troisième sous-couche de la couche O, réalisant ainsi le lanthane (figure 6.3). Ensuite, un électron vient enfin se loger dans la sous-couche la plus extérieure de la couche N. Un par un, les éléments des terres rares ajoutent des électrons à la couche N, jusqu'à l'élément 71, le lutécium, qui finalement la complète. Les électrons du lutécium se répartissent ainsi : deux dans la couche K, huit dans la couche L, dix-huit dans la couche M, trente-deux ans la couche N, neuf dans la couche D (deux sous-couches pleines plus un électron dans la sous-couche suivante), et deux dans la couche P (sous-couche la plus intérieure pleine).

Nous commençons enfin à voir pourquoi les terres rares, et quelques autres groupes d'éléments de transition, sont si semblables. Le caractère de différenciation des éléments, décisif pour leurs propriétés chimiques, est la configuration des électrons dans leur couche extérieure. Le carbone par exemple, avec quatre électrons dans la couche externe, et l'azote, avec cinq, ont des propriétés totalement différentes. En revanche, dans les séries où les électrons se consacrent au remplissage de sous-couches intérieures sans toucher à la couche externe, on trouve moins de variété dans les propriétés. C'est ainsi que le fer, le cobalt et le nickel (26, 27 et 28), ayant

Figure 6.3. Représentation schématique de l'empiétement des couches et sous-couches électroniques dans le lanthane. La sous-couche la plus extérieure de la couche N reste à pourvoir.



la même configuration d'électrons sur la couche externe – une sous-couche N remplie de deux électrons – ont des comportements chimiques assez voisins. Leurs différences électroniques internes (dans une sous-couche M) sont bien dissimulées sous leur surface électronique similaire. Il en va de même, a fortiori, pour les éléments de la série des terres rares. Leurs différences (dans la couche N) sont enfouies sous non pas une, mais deux configurations électroniques extérieures (dans les couches O et P) qui sont identiques pour tous ces éléments. Rien d'étonnant à ce que, chimiquement, ces éléments se ressemblent comme des gouttes d'eau.

Le peu d'applications des métaux des terres rares, et la difficulté de les séparer, expliquent la pauvreté des efforts que les chimistes ont consacrés à cette tâche – tout au moins jusqu'à la fission de l'atome d'uranium. La séparation devint alors une nécessité urgente car la poursuite du projet de bombe nucléaire exigeait la séparation et l'identification, rapides et sans ambiguïté, des isotopes radioactifs de certains des éléments qui figuraient parmi les produits principaux de la fission.

Ce problème fut bientôt résolu en ayant recours à une technique chimique qui avait été introduite en 1906 par le botaniste russe Mikhail Semenovitch Tswett, et qu'il avait baptisée *chromatographie* (« écriture en couleur »). Tswett avait découvert qu'il pouvait séparer des pigments de plantes, chimiquement très semblables, en en égouttant une solution dans une colonne de craie pulvérisée. Il dissolvait sa mixture de pigments végétaux dans un éther de pétrole qu'il versait sur la craie. Puis il versait du solvant pur. En s'écoulant doucement à travers la poudre de craie, chaque pigment avait une vitesse différente à cause des variations des forces d'adhérence des pigments à la poudre selon leur nature. Il en résultait une séparation des pigments en une série de bandes diversement colorées. Soumises à un lavage continu, les différentes substances s'égouttaient successivement au bas de la colonne.

Le monde scientifique négligea longtemps la découverte de Tswett, probablement parce qu'il n'était qu'un botaniste... russe de surcroît, alors que les chefs de file de la séparation des substances inséparables étaient, à cette époque, des biochimistes allemands. Mais le procédé fut popularisé après sa redécouverte en 1931 par Richard Willstätter, un biochimiste allemand. (Willstätter était titulaire du prix Nobel de chimie 1915, pour ses remarquables travaux sur les pigments végétaux. Quant à Tswett, il est, pour autant que je sache, resté oublié.)

On s'est aperçu que la chromatographie à travers une colonne de poudre convenait pour toutes sortes de mélanges, aussi bien incolores que multicolores. L'oxyde d'aluminium et l'amidon se sont avérés supérieurs à la craie pour séparer les molécules ordinaires. Lorsqu'on sépare des ions, le procédé s'appelle *échange d'ions,* et les premiers agents efficaces dans ce rôle ont été des composés baptisés *zéolites*. On peut ôter, par exemple, les ions calcium et magnésium d'une *eau dure* en faisant couler cette eau à travers une colonne de zéolite. Les ions calcium et magnésium s'attachent à la zéolite tandis que les ions sodium, primitivement dans la zéolite, prennent leur place dans la solution, et c'est ainsi de *l'eau douce* qui s'égoutte du bas de la colonne. Il faut de temps en temps refaire le plein en ions sodium en versant une solution concentrée de sel (chlorure de sodium). La mise au point des *résines échangeuses d'ions* en 1935 a constitué une amélioration. Ces substances synthétiques peuvent être conçues pour une tâche spécifique. Certaines résines, par exemple, substituent des ions

hydrogène à des ions positifs, tandis que d'autres remplacent des ions négatifs par des ions hydroxyles ; une combinaison des deux permet de retenir la plupart des sels de l'eau de mer. Des systèmes contenant ces résines faisaient partie des nécessaires de survie à bord des canots de sauvetage pendant le dernière guerre.

C'est le chimiste américain Frank Harold Spedding qui a adapté la chromatographie par échange d'ions à la séparation des terres rares. Il a découvert que les éléments sortaient d'une colonne d'échange d'ions en ordre inverse de leurs numéros atomiques, ce qui permettait de les séparer rapidement aussi bien que de les identifier. C'est en fait de cette façon que la découverte du prométhéum, l'élément manquant 61, fut confirmée à partir des petites quantités trouvées dans les produits de fission.

On peut maintenant, grâce à la chromatographie, préparer en grande quantité (des kilos, ou même des tonnes) des éléments purifiés du groupe des terres rares. Il se trouve que les terres rares ne sont finalement pas si rares : les plus rares d'entre elles (à l'exception du prométhéum) sont plus communes que l'or et l'argent, et les plus courantes – le lanthane, le cérium et le néodyme – sont plus abondantes que le plomb. Au total, les métaux des terres rares participent à la composition de l'écorce terrestre dans une plus grande proportion que le cuivre et l'étain ensemble. Les scientifiques ont donc a peu près abandonné le terme de *terres rares* et désignent plutôt cette série d'éléments sous le nom de *lanthanides*, d'après leur tête de liste. Les éléments des lanthanides n'avaient à vrai dire guère été utilisés individuellement, mais la possibilité de les séparer a multiplié leurs utilisations, et on en consomme maintenant une dizaine de milliers de tonnes par an. Le mischmétal est un mélange composé principalement de cérium, de lanthane et de néodyme, qui entre dans la constitution des pierres à briquets pour les trois quarts de leur poids. On utilise un mélange d'oxydes de lanthanides pour polir le verre, tandis que ces oxydes, ajoutés au verre, lui confèrent les propriétés souhaitées. On se sert de mélanges d'oxydes d'europium et d'yttrium pour obtenir la phosphorescence rouge sur les tubes de télévision en couleur, etc.

LES ACTINIDES

Mais les bénéfices d'une meilleure compréhension des lanthanides ne se bornent pas à leurs applications pratiques. Ces nouvelles connaissances ont apporté une clé pour la chimie des éléments de la fin de la classification périodique, y compris les éléments de synthèse.

La série des éléments lourds en question commence avec l'actinium (89). Celui-ci se trouve dans le tableau juste au-dessous du lanthane. L'actinium a deux électrons dans la couche Q, tout comme le lanthane dans la couche P. Le 89e et dernier électron de l'actinium vient se loger dans la couche P, comme le 57e lanthane se met dans la couche O. La question est alors la suivante : les éléments à la suite de l'actinium continuent-ils à ajouter des électrons à la couche P pour rester des éléments de transition ordinaires, ou alors, ne suivraient-ils pas l'exemple des éléments qui suivent le lanthane, dont les électrons viennent combler la sous-couche inférieure oubliée ? Dans ce cas, l'actinium pourrait donner le départ à une nouvelle série de métaux du type terres rares que l'on baptiserait naturellement *actinides*.

Les éléments naturels dans cette supposée série des actinides sont l'actinium, le thorium, le protoactinium et l'uranium. Ces éléments n'avaient guère été étudiés avant 1940. Le peu que l'on connaissait de leur chimie indiquait qu'il s'agissait d'éléments de transition ordinaires. Mais après l'addition à cette liste et l'étude intensive du neptunium et du plutonium, il s'avéra que ces éléments artificiels manifestaient une étroite ressemblance chimique avec l'uranium. C'est ce qui poussa Glenn Seaborg à suggérer que les éléments lourds suivaient en fait le schéma des lanthanides et remplissaient la quatrième sous-couche, enfouie mais vacante, de la couche O. Cette sous-couche est comblée avec le lawrencium et il existe donc quinze actinides, en parfaite analogie avec les quinze lanthanides. On en a une confirmation importante avec le fait que la chromatographie par échange d'ions sépare les actinides de la même façon que les lanthanides.

Les éléments 104 (le rutherfordium) et 105 (le hahnium) sont des *transactinides* qui, les chimistes en sont pratiquement certains, viennent se placer sous l'hafnium et le tantale, les deux éléments qui suivent les lanthanides.

Les gaz

Depuis les débuts de la chimie on s'est aperçu que de nombreuses substances pouvaient exister sous forme de gaz, de liquide ou de solide, selon la température. L'eau en est l'exemple le plus ordinaire : suffisamment refroidie, elle devient de la glace solide ; suffisamment chauffée, elle se change en vapeur gazeuse. Van Helmont, le premier utilisateur du mot *gaz,* distinguait les corps qui sont gazeux à température ordinaire, comme le dioxyde de carbone, de ceux qui, comme la vapeur, ne sont gazeux qu'à température élevée. Il appelait ces derniers des *vapeurs,* et l'on parle encore de *vapeurs d'eau* plutôt que de *gaz aqueux* ou d'*eau gazeuse.*

L'étude des gaz, ou vapeurs, en partie parce qu'ils se prêtent volontiers à l'analyse quantitative, n'a cessé de fasciner les chimistes. Les règles de leur comportement sont plus simples et plus faciles à établir que celles qui régissent les liquides et les solides.

LA LIQUÉFACTION

Le physicien français Jacques Alexandre César Charles découvrit en 1787 que lorsqu'on refroidissait un gaz, chaque degré de refroidissement supplémentaire diminuait le volume de ce gaz d'environ 1/273 de son volume à 0 °C ; et inversement, toute élévation de température de un degré augmentait son volume de 1/273. L'expansion sous l'influence de la chaleur ne posait pas de difficulté logique, mais si la contraction avec le froid devait se poursuivre selon la *loi de Charles* (comme on l'appelle encore maintenant), un gaz aurait dû être réduit à rien en atteignant la température de –273 °C ! Le paradoxe ne gênait pas trop les chimistes, sûrs qu'ils étaient de l'impossibilité pour la loi de Charles de rester valide en

descendant l'échelle des températures, car tout gaz devait finir par se condenser en liquide, et les liquides sont loin de se contracter comme des gaz lorsque la température baisse. Il n'en restait pas moins que les chimistes n'avaient au début aucun procédé pour atteindre les très basses températures qui leur permettraient de voir effectivement ce qui se passait.

L'avènement de la théorie atomique, dans laquelle les gaz sont des ensembles de molécules, présentait la situation en des termes nouveaux. On réalisait désormais que le volume dépendait de la vitesse des molécules. Plus la température était élevée, plus les molécules allaient vite, plus elles exigeaient des « coudées franches » et plus le volume était grand. Inversement, sous une température plus basse, elles se déplaçaient plus lentement, exigeaient moins d'espace, et le volume était plus petit. Le physicien britannique William Thomson, alors qu'il venait juste d'être élevé au rang de pair du royaume, avec le titre de Lord Kelvin, émit l'idée que c'était l'énergie moyenne des molécules qui s'abaissait de 1/273 à chaque degré. Alors que l'on ne pouvait attendre du volume qu'il disparaisse complètement, c'était chose possible pour l'énergie. Thomson prétendait qu'à –273 °C l'énergie devait s'annuler. Ainsi, cette valeur de –273 °C devait représenter la plus basse température possible. Cette température (maintenant de –273,16 °C, selon les méthodes de mesure modernes) devait donc être le *zéro absolu* ou, comme on dit parfois, le *zéro Kelvin.* Sur cette échelle absolue, le point de fusion de la glace est à 273 °K. (Voir la figure 6.4 pour les échelles Fahrenheit, Celsius et Kelvin.)

Selon ce point de vue, il était encore plus certain que les gaz devaient tous se liquéfier en approchant du zéro absolu. L'énergie disponible diminuant, les molécules de gaz devaient exiger si peu d'espace qu'elles se resserraient et entraient en contact les unes avec les autres. En d'autres termes elles constituaient des liquides, car les propriétés des liquides s'expliquent en les supposant formés de molécules en contact, mais qui ont encore assez d'énergie pour glisser librement les unes par rapport aux autres. C'est pour cette raison que les liquides s'écoulent et qu'ils s'adaptent sans difficulté à la forme du récipient qui les contient.

Comme l'énergie décroît toujours, à mesure que le température baisse, il en reste finalement trop peu pour que les molécules puissent continuer à passer leur chemin, et elles occupent alors des positions fixées autour desquelles elles peuvent osciller, mais qu'elles ne peuvent quitter. En d'autres termes, le liquide est gelé et devient un solide. Il semblait donc pour Kelvin qu'en approchant du zéro absolu, les gaz devaient non seulement se liquéfier, mais se solidifier.

Les chimistes souhaitaient naturellement établir l'exactitude des prédictions de Kelvin en abaissant d'abord la température au point où tous les gaz seraient liquides, puis géleraient, avant de parvenir effectivement au zéro absolu. (Il y a toujours, dans tout horizon lointain, un appel à la conquête.)

Les scientifiques avaient exploré les froids extrêmes avant que Kelvin ait défini le but ultime. Michael Faraday avait découvert que l'on pouvait liquéfier certains gaz en les comprimant, même à température ordinaire. Il se servait d'un tube de verre épais, courbé comme un boomerang. Il plaçait au fond du tube la substance qui devait fournir le gaz qui l'intéressait. Il scellait ensuite l'ouverture du tube. Il plongeait le fond du tube dans l'eau chaude, produisant ainsi du gaz en quantité toujours plus importante ; comme le gaz était confiné dans le tube, il exerçait une

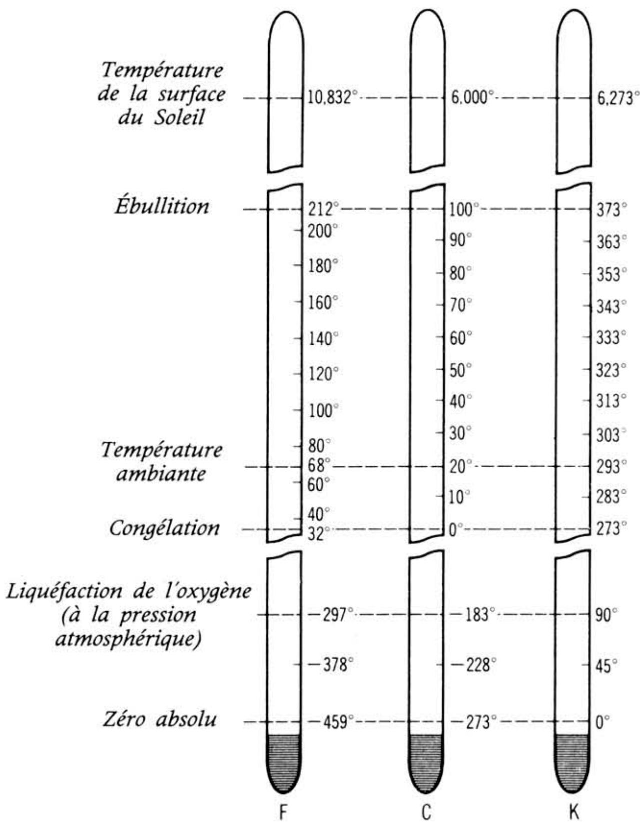


Figure 6.4. Comparaison des échelles thermométriques Fahrenheit, Celsius (ou Centigrade) et Kelvin.

pression de plus en plus grande. Faraday maintenait l'autre bout du tube dans un récipient de glace pilée. Le gaz contenu dans cette extrémité, étant soumis en même temps à une haute pression et à une basse température, avait une chance de se liquéfier. Faraday parvint de cette façon à liquéfier le chlore gazeux en 1823 alors que, sous pression normale, le point de liquéfaction du chlore est à −34,5 °C (238,7 °K).

Un chimiste français, C.S.A. Thilorier, utilisa en 1835 le procédé de Faraday pour liquéfier le dioxyde de carbone CO_2 sous pression, en employant des cylindres métalliques capables de résister à de plus hautes pressions que les tubes de verre. Après avoir liquéfié une quantité considérable de dioxyde de carbone, il laissait celui-ci s'échapper du tube par une petite ouverture.

Le dioxyde de carbone liquide, alors exposé à la température ambiante, devait naturellement s'évaporer rapidement. Lorsqu'un liquide s'évapore, ses molécules sont écartelées de leurs voisines et retrouvent leur individualité pour se mouvoir en toute liberté. Dans un liquide, il y a attraction mutuelle entre les molécules, et échapper à cette attraction

requiert de l'énergie. En cas d'évaporation rapide, l'énergie n'a pas le temps de parvenir en quantité suffisante dans le système (sous forme de chaleur), et il ne reste comme seule source d'énergie, pour alimenter l'évaporation, que le liquide lui-même. Lorsqu'un liquide s'évapore rapidement, le liquide résiduel voit donc sa température s'abaisser.

(Nous expérimentons nous-mêmes ce phénomène, car le corps humain transpire toujours un peu, et l'évaporation de la fine couche d'eau sur notre peau emporte la chaleur de celle-ci et nous refroidit. Plus il fait chaud, plus nous transpirons ; et si l'humidité de l'air est telle que l'évaporation ne puisse avoir lieu, la sueur s'accumule sur notre corps et la situation devient très désagréable. L'exercice physique, en stimulant les réactions productrices de chaleur dans notre corps, accroît également la transpiration, et nous nous retrouvons alors dans une situation aussi inconfortable que lorsque l'atmosphère est humide.)

Lorsque Thilorier laissait le dioxyde de carbone liquide s'évaporer, la température du liquide s'abaissait à mesure de l'évaporation, jusqu'à ce que le dioxyde de carbone gèle. On avait, pour la première fois, formé du dioxyde de carbone solide.

Le dioxyde de carbone liquide ne reste stable que sous pression. A la pression ordinaire, le dioxyde de carbone solide *sublime* – c'est-à-dire qu'il s'évapore directement en gaz carbonique, sans fondre d'abord. Le point de sublimation du dioxyde de carbone solide est à –78,5 °C (194,7 °K).

Le dioxyde de carbone solide a un aspect de glace laiteuse (mais il est beaucoup plus froid) ; on l'appelle *glace carbonique* ou, parce qu'il ne produit pas de liquide, *glace sèche* chez les Anglo-Saxons. On en produit environ 400 000 tonnes par an, surtout pour la conservation des aliments réfrigérés.

Le refroidissement par évaporation a constitué une révolution dans les activités humaines. Avant le XIXe siècle on utilisait la glace, si on en avait, pour conserver les aliments. La glace pouvait être emmagasinée en hiver et conservée, bien isolée, jusqu'en été ; ou elle pouvait être apportée des régions montagneuses. C'était au mieux un procédé pénible et fastidieux, et la plupart des gens devaient s'accommoder de la chaleur de l'été (quand ce n'était pas toute l'année).

Dès 1755, le chimiste écossais William Cullen produisit de la glace en faisant le vide au-dessus d'une grande masse d'eau, provoquant une évaporation rapide qui refroidissait l'eau jusqu'au point de congélation. Ce procédé n'entrait toutefois pas en concurrence avec la glace naturelle ; pas plus qu'il n'était utilisable indirectement pour refroidir les denrées alimentaires, car la formation de glace obstruait les tuyaux.

De nos jours, on liquéfie avec un compresseur un gaz convenablement choisi que l'on laisse revenir à la température ambiante. On fait ensuite circuler le liquide dans un serpentin autour de l'enceinte dans laquelle se trouvent les aliments. L'évaporation du liquide absorbe la chaleur de l'enceinte. Le gaz qui ressort est à nouveau liquéfié par le compresseur, refroidi et recyclé. Le processus est continu, et on pompe ainsi la chaleur de l'enceinte fermée vers l'atmosphère extérieure. C'est le *réfrigérateur*, qui a remplacé l'ancienne *glacière*.

En 1834, un inventeur américain, Jacob Perkins, breveta en Angleterre l'utilisation de l'éther en guise de réfrigérant. D'autres gaz comme l'ammoniac et le dioxyde de soufre ont été utilisés. Tous ces réfrigérants avaient l'inconvénient d'être toxiques ou inflammables. Mais en 1930, le chimiste américain Thomas Midgley a découvert le dichlorodifluorométhane

thane (CF_2Cl_2), plus connu sous la désignation commerciale de Fréon. C'est un gaz inoffensif (comme Midgley le démontra en en inhalant en public) et ininflammable, qui convient parfaitement. Avec le fréon, la réfrigération à domicile devenait accessible.

(Bien que l'on ait prouvé la totale inocuité du fréon et autres *hydrocarbures fluorés* pour l'organisme humain, on a commencé à avoir des doutes ces dernières années en ce qui concerne leur effet sur la couche d'ozone de la haute atmosphère, comme je l'ai dit dans le chapitre précédent.)

La réfrigération modérée de grandes quantités d'air s'appelle la *climatisation ;* l'air y est aussi filtré et débarrassé d'une partie de son humidité. Le premier appareil de climatisation pratique fut mis au point en 1902 par l'inventeur américain Willis Haviland Carrier ; depuis la Deuxième Guerre mondiale, la climatisation est devenue quasi universelle... tout au moins dans les grandes villes d'Amérique du Nord.

Thilorier, encore lui, ajouta du dioxyde de carbone solide dans un liquide appelé *éther* (plus connu de nos jours comme anesthésique ; voir le chapitre 11). L'éther bout à basse température et s'évapore rapidement. Ceci, joint à la basse température du dioxyde de carbone solide en cours de sublimation, fit descendre la température jusqu'à −110 °C (163,2 °K).

Faraday revint, en 1845, au problème de la liquéfaction des gaz sous l'effet combiné d'une basse température et d'une haute pression, au moyen d'un mélange réfrigérant de dioxyde de carbone solide et d'éther. Même avec ce mélange et des pressions plus élevées qu'auparavant, il restait six gaz qu'il ne parvint pas à liquéfier. Il s'agissait de l'hydrogène, de l'oxygène, de l'azote, du monoxyde de carbone, du dioxyde d'azote et du méthane, qu'il désigna sous le nom de *gaz permanents.* A cette liste, nous pourrions ajouter cinq autres gaz ignorés de Faraday, le fluor et les quatre gaz rares : l'hélium, le néon, l'argon et le krypton.

Mais en 1869, le physicien irlandais Thomas Andrews déduisit de ses expériences qu'il existait pour chaque gaz une *température critique* au-dessus de laquelle la liquéfaction était impossible, quelle que fût la pression. Cette hypothèse fut par la suite justifiée théoriquement par le physicien hollandais Johannes Diderik Van der Waals, qui se vit pour cela décerner le prix Nobel de physique en 1910.

Pour liquéfier un gaz donné il fallait donc être certain de travailler à une température inférieure à la température critique, sinon on agissait en pure perte. On ne ménagea pas les efforts pour atteindre des températures encore plus basses qui permettraient de liquéfier les gaz les plus rétifs. L'astuce consistait à utiliser une méthode de refroidissements successifs. On utilisa d'abord du dioxyde de soufre liquide, refroidi par évaporation, pour liquéfier le dioxyde de carbone ; ensuite le dioxyde de carbone pour liquéfier un gaz plus rétif ; et ainsi de suite. Le physicien suisse Raoul Pictet parvint finalement en 1877 à liquéfier de l'oxygène à la température de −140 °C (133 °K) et sous une pression de 500 atmosphères (500 kilos par centimètre carré). A peu près à la même époque, le physicien français Louis Paul Cailletet put liquéfier aussi l'azote et le monoxyde de carbone. Naturellement, ces liquides permirent immédiatement d'atteindre des températures encore plus basses. On a finalement trouvé qu'à pression ordinaire, les points de liquéfaction de l'oxygène, du monoxyde de carbone et de l'azote étaient respectivement à −183 °C (90 °K), −190 °C (83 °K) et −195 °C (78 °K).

En 1895, l'ingénieur chimiste anglais William Hampson et le physicien allemand Karl von Linde ont mis au point, indépendamment, une méthode pour liquéfier de grandes quantités d'air. L'air était d'abord comprimé, puis refroidi à température ordinaire. Une détente permettait ensuite à cet air de se refroidir considérablement. On utilisait cet air froid pour refroidir un réservoir d'air comprimé, jusqu'à ce que cet air lui aussi soit froid. L'expansion de cet air comprimé lui permettait ensuite de devenir encore beaucoup plus froid. Le processus était répété jusqu'à ce que l'air, devenant de plus en plus froid, finisse par se liquéfier.

L'air liquide, abondant et bon marché, est facilement séparé en oxygène liquide et en azote liquide. L'oxygène a pu ainsi être utilisé pour les chalumeaux et les applications médicales, et l'azote pour les applications dans lesquelles son inertie le rend souhaitable. Ainsi, les lampes incandescentes, remplies d'azote, voient leur filament chauffé à blanc supporter cette haute température plus longtemps (avant de se casser du fait de la lente évaporation du métal) que si ce même filament brillait dans une ampoule vide. L'air liquide est également une source de composants secondaires comme l'argon et les autres gaz rares.

L'hydrogène résista jusqu'en 1900 à toutes les tentatives de liquéfaction. C'est le chimiste écossais James Dewar qui y parvint finalement en introduisant un nouveau stratagème. Lord Kelvin (William Thomson) et le physicien anglais James Prescott Joule avaient montré que, même à l'état gazeux, on pouvait facilement refroidir un gaz en le laissant se dilater tout en empêchant la chaleur extérieure de parvenir à ce gaz, à condition que la température de départ soit déjà assez basse. Dewar avait donc refroidi de l'hydrogène comprimé jusqu'à –200 °C, dans un réservoir entouré d'azote liquide, et par une expansion de cet hydrogène superfroid l'avait encore refroidi, puis il avait répété ce cycle en faisant repasser l'hydrogène de plus en plus froid dans la tuyauterie. L'hydrogène comprimé, soumis à cet *effet Joule-Thomson*, devint finalement liquide à la température de –240 °C (33 °K). A des températures encore plus basses, Dewar parvint à obtenir de l'hydrogène solide.

Pour conserver ses liquides superfroids, Dewar conçut des récipients spéciaux à doubles parois en verre argenté entre lesquelles on faisait le vide. Dans le vide, la chaleur ne peut être échangée que par rayonnement, qui est un mécanisme relativement lent, et le revêtement d'argent réfléchit le rayonnement incident. Ces *vases de Dewar* sont les ancêtres directs de la bouteille Thermos, si courante à présent.

LES PROPERGOLS

Les gaz liquéfiés se sont trouvés, avec le développement des fusées, promus à des hauteurs astronomiques. Il faut pour les fusées des réactions chimiques extrêmement rapides, produisant de grandes quantités d'énergie. Le produit le plus pratique est une combinaison de combustible liquide, comme l'alcool ou le kérosène, et d'oxygène liquide. Dans tous les cas, la fusée doit transporter son oxygène, ou un autre agent oxydant, car elle ne dispose d'aucun approvisionnement naturel en oxygène lorsqu'elle quitte l'atmosphère. L'oxygène doit être sous forme liquide, car les liquides sont plus denses que les gaz, ce qui permet de stocker dans les réservoirs plus d'oxygène à l'état liquide que sous la forme gazeuse. C'est pourquoi

l'oxygène liquide est devenu un produit très recherché dans l'industrie des fusées.

On mesure l'efficacité d'un mélange de carburant et d'oxydant par une grandeur appelée *poussée spécifique*. C'est la force de poussée produite par la combustion d'une unité de poids du mélange en une seconde. Pour un mélange de kérosène et d'oxygène, cette poussée spécifique vaut 242. Comme la charge utile que peut emporter une fusée dépend de la poussée spécifique, la recherche de combinaisons plus efficaces a été intense. De ce point de vue, le meilleur combustible est l'hydrogène liquide. En se combinant avec l'oxygène liquide il peut produire une poussée spécifique de l'ordre de 350. Si l'on pouvait utiliser de l'ozone ou du fluor liquides, au lieu d'oxygène, la poussée spécifique s'élèverait jusqu'à 370.

Des métaux légers, comme le lithium, le bore, le magnésium, l'aluminium et surtout le béryllium, fournissent en se combinant avec l'oxygène encore plus d'énergie que l'hydrogène. Mais certains sont rares, et leur combustion est toujours hérissée de difficultés techniques, venant des fumées, des dépôts d'oxydes, etc.

Il existe aussi des combustibles solides qui sont leur propre oxydant (comme la poudre à canon, le premier combustible de fusée, mais beaucoup plus efficaces). De tels combustibles sont appelés *monopropergols*, car il n'ont pas besoin d'oxydant séparé et constituent par eux-mêmes tout le matériau de propulsion, le *propergol*. Un propergol qui exige un approvisionnement en oxydant séparé est un *bipropergol*. Les monopropergols sont de stockage et de manipulation simples, et ils brûlent rapidement mais de façon contrôlée. La difficulté consiste à mettre au point un monopropergol dont la poussée spécifique approcherait celle des bipropergols.

L'hydrogène atomique, que Langmuir avait déjà utilisé pour son chalumeau, constitue une autre possibilité. On a calculé qu'un moteur de fusée fonctionnant par recombinaison des atomes d'hydrogène en molécules pourrait produire une poussée spécifique supérieure à 1 300. Le problème réside dans le stockage de l'hydrogène atomique. La meilleure voie semble être pour l'instant dans un refroidissement important et rapide des atomes immédiatement après leur formation ; les recherches menées au Bureau des mesures aux États-Unis indiquent que l'on peut conserver les atomes d'hydrogène libres en les piégeant dans un corps solide à très basse température, comme l'oxygène ou l'argon gelés. S'il suffisait de presser un bouton, pour ainsi dire, pour que les gaz gelés commencent à se réchauffer et s'évaporer, les atomes d'hydrogène seraient alors libérés et pourraient se recombiner. Avec un tel solide, même s'il ne permettait de retenir que 10 % de son poids en atomes d'hydrogène libres, on disposerait d'un propergol meilleur que tous ceux qui existent à l'heure actuelle. Mais il faudrait naturellement que la température soit très basse, beaucoup plus basse que celle de l'hydrogène liquide. Ces solides devraient être maintenus à environ –272 °C, ou juste un degré au-dessus du zéro absolu.

Dans un autre ordre d'idées, on a la possibilité d'éjecter des ions, plutôt que les gaz d'échappement du propergol brûlé. Chaque ion, de masse minuscule, donnerait une toute petite impulsion, mais le processus pourrait être poursuivi lontemps. Un véhicule placé sur orbite au moyen de la poussée élevée, mais de courte durée, d'un propergol chimique pourrait ensuite accélérer doucement, libre de tout frottement dans

l'espace, sous l'effet permanent d'ions éjectés à une vitesse voisine de celle de la lumière. Le césium est le corps qui convient le mieux pour la propulsion ionique, car c'est l'élément auquel il est le plus facile d'arracher des électrons pour former un ion césium. Il n'y a plus qu'à établir un champ électrique pour accélérer l'ion césium et l'expulser par la tuyère de la fusée.

SUPRACONDUCTEURS ET SUPERFLUIDES

Pour revenir au domaine des basses températures, même la liquéfaction et la solidification de l'hydrogène n'en constituent pas l'ultime prouesse. Au moment où l'hydrogène cédait, les gaz rares étaient découverts ; le plus léger de ceux-ci, l'hélium, refusait obstinément de se liquéfier aux températures les plus basses alors atteintes. Le physicien hollandais Heike Kammerlingh Onnes réussit en 1908 à triompher de l'hélium en ajoutant une étape supplémentaire au procédé de Dewar. Avec l'hydrogène liquide, il refroidissait jusqu'à –255 °C (18 °K) de l'hélium gazeux sous pression, puis une expansion permettait à ce gaz de se refroidir encore plus. Il parvenait ainsi à liquéfier le gaz. Plus tard, par évaporation de l'hélium liquide, il atteignit la température de liquéfaction de l'hélium à la pression atmosphérique (4,2 °K), température à laquelle *tous* les autres corps sont solides, et même des températures allant jusqu'à 0,7 °K. Onnes fut récompensé par le prix Nobel de physique en 1913 pour ses travaux dans le domaine des basses températures. (De nos jours, la liquéfaction de l'hélium n'a rien de compliqué. Le chimiste américain Samuel Cornette Collins a inventé en 1947 un *cryostat* qui, par une série de compressions et expansions, peut produire quelques litres d'hélium liquide à l'heure.)

Mais Onnes ne s'est pas contenté d'atteindre des températures records. Il a été le premier à montrer qu'à ces basses températures la matière manifeste des propriétés originales. Parmi ces propriétés il y a l'étrange phénomène baptisé *supraconductivité*. En 1911, Onnes étudiait la résistance électrique du mercure à basse température. On pensait alors que la résistance devait décroître régulièrement à mesure que l'extraction de chaleur diminuait l'amplitude des vibrations des atomes du métal. Mais à 4,12 °K, la résistance du mercure s'annula soudainement ! Un courant électrique pouvait alors le traverser sans aucune perte. On s'est rapidement aperçu que l'on pouvait rendre supraconducteurs d'autres métaux. Le plomb, par exemple, devient supraconducteur à 7,22 °K. Un courant électrique de plusieurs centaines d'ampères, lancé dans un anneau de plomb maintenu à cette température par de l'hélium liquide, peut continuer à circuler dans l'anneau pendant deux ans et demi, sans que l'on puisse déceler une quelconque décroissance de l'intensité.

A mesure que l'on atteignait des températures de plus en plus basses, d'autres métaux sont venus s'ajouter à la liste des corps supraconducteurs. L'étain devient supraconducteur à 3,73 °K ; l'aluminium à 1,20 °K ; l'uranium à 0,8 °K ; le titane à 0,53 °K ; l'hafnium à 0,35 °K. (On connaît maintenant environ 1 400 éléments et alliages différents qui manifestent une supraconductivité.) Mais le fer, le nickel, le cuivre, l'or, le sodium et le potassium doivent avoir des points de transition encore plus bas – si tant est qu'ils puissent être supraconducteurs – car la température la plus basse atteinte actuellement n'a pu les amener à cet état. Le point de transition le plus élevé trouvé pour un élément métallique est celui du technétium, qui devient supraconducteur au-dessous de 11,2 °K.

Par immersion dans un liquide dont le point d'ébullition est assez bas, on peut facilement maintenir un corps à cette température. Pour parvenir à une température plus basse, il faut faire appel à un liquide dont le point d'ébullition soit encore plus bas. L'hydrogène liquide bout à 20,4 °K, et il serait bien commode de disposer d'un corps supraconducteur dont la température de transition soit au moins aussi élevée. On pourrait étudier la supraconductivité de systèmes refroidis par l'hydrogène liquide. A défaut, on est obligé de recourir au seul liquide dont le point d'ébullition soit inférieur, à savoir l'hélium, plus rare, plus cher et plus difficile à manipuler. Il existe quelques alliages, en particulier avec le niobium métallique, dont les températures de transition sont plus élevées que celles des métaux purs. Finalement, on a découvert en 1968 un alliage de niobium, d'aluminium et de germanium qui reste supraconducteur à 21 °K. La supraconductivité aux températures de l'hydrogène liquide devient réalisable, mais il s'en faut de peu.

On pense immédiatement à une application pratique de la supraconductivité dans le domaine du magnétisme. En circulant dans un bobinage autour d'un noyau de fer, un courant électrique peut produire un fort champ magnétique : plus le courant est intense, plus le champ est fort. Malheureusement aussi, dans les circonstances ordinaires, plus le courant est intense et plus grande est la chaleur dégagée ; il y a donc une limite à ce qui est réalisable. Par contre, dans les fils supraconducteurs, l'électricité circule sans produire de chaleur ; tout semblerait donc indiquer que l'on puisse accumuler de plus en plus de courant dans ces fils, pour réaliser des *électro-aimants* qui ne consommeraient qu'une fraction de la puissance requise dans des conditions ordinaires. Mais il y a un problème.

En même temps que la supraconductivité, se manifeste une autre propriété qui a trait au magnétisme. Dès qu'un corps passe à l'état supraconducteur, il devient également parfaitement *diamagnétique,* c'est-à-dire qu'il expulse les lignes de champ magnétique. Ce phénomène est appelé *effet Meissner*, d'après le physicien allemand Walther Meissner qui l'a découvert en 1933.

Mais en augmentant le champ magnétique, on peut faire disparaître la supraconductivité du corps, et du même coup anéantir tout espoir de supermagnétisme, même à des températures bien inférieures au point de transition. Tout se passe comme si une partie des lignes de champ parvenaient à pénétrer le corps, lorsqu'on en accumule suffisamment autour,et éliminaient alors toute supraconductivité.

On a tenté de trouver des corps supraconducteurs qui tolèrent des champs magnétiques élevés. Un alliage d'étain et de zirconium, par exemple, a une température de transition, relativement élevée, de 18 °K. Il accepte un champ magnétique de 25 Tesla, ce qui est vraiment beaucoup. On avait découvert ce fait en 1954, mais il a fallu attendre 1960 pour disposer d'une technique d'étirement des fils de cet alliage très fragile. On peut encore faire mieux avec un mélange de vanadium et de gallium : on a réalisé des électro-aimants supraconducteurs dont le champ atteint 50 Tesla.

C'est dans l'hélium lui-même que l'on a découvert un autre phénomène étonnant, caractéristique des basses températures. Il s'agit de la *superfluidité.*

L'hélium est la seule substance connue qu'il soit impossible de solidifier, même au zéro absolu. Il contient toujours un peu d'énergie résiduelle, que l'on ne peut extraire même au zéro absolu (de sorte qu'à toutes fins

pratiques le contenu énergétique est « effectivement nul »), mais suffisante pour que les atomes d'hélium, particulièrement peu attirés les uns par les autres, gardent leur liberté et restent donc liquides. Le physicien allemand Hermann Walther Nernst avait en fait montré, en 1905, que ce n'était pas l'énergie d'un corps qui s'annulait au zéro absolu, mais une grandeur reliée à celle-ci, *l'entropie*. Le prix Nobel de chimie lui fut décerné pour cela en 1920. Mais attention, je n'ai pas dit que l'hélium solide n'existe pas ; il a été produit en 1926, à des températures inférieures à 1 °K, sous une pression d'environ 25 atmosphères.

En 1935, Wilhem Hendrik Keesom, qui avait réussi à solidifier l'hélium, et sa sœur A.P. Keesom, tous deux membres du laboratoire d'Onnes, s'aperçurent que l'hélium liquide en dessous de 2,2 °K se mettait à conduire la chaleur presque parfaitement. Il conduit la chaleur si rapidement – à la vitesse du son, en fait – que toutes les parties de l'hélium sont toujours à la même température. Il ne peut bouillir – comme le ferait tout liquide ordinaire grâce à des zones localisées plus chaudes où se forment des bulles de vapeur – car il n'y a aucune zone localement plus chaude dans l'hélium liquide (si tant est que l'on puisse parler de zones chaudes à propos d'un liquide au-dessous de 2 °K). Lorsque le liquide s'évapore, la surface supérieure du liquide disparaît tranquillement – elle se décolle par plaques, pour ainsi dire.

Le physicien russe Piotr Leonidovitch Kapitsa a étudié cette propriété et trouvé que la raison pour laquelle l'hélium conduit si bien la chaleur vient du fait qu'il s'écoule particulièrement facilement, transportant presque instantanément la chaleur d'un point à un autre, au moins deux cents fois plus rapidement que le cuivre, le meilleur conducteur thermique après l'hélium. Il s'écoule encore plus facilement qu'un gaz ; sa viscosité n'est que le millième de celle de l'hydrogène gazeux, et il peut fuir à travers une ouverture si petite qu'elle bloquerait un gaz. De plus, le liquide superfluide forme une pellicule sur le verre qui se meut aussi rapidement que si le liquide s'écoulait par un trou. Si un récipient ouvert contenant le liquide est placé dans un récipient plus grand où le même liquide se trouve à un niveau inférieur, le fluide va grimper sur la paroi de verre et passer par-dessus le rebord du récipient jusqu'à égalisation des niveaux.

L'hélium est le seul corps qui présente cette propriété de superfluidité. Le superfluide a un comportement si différent de l'hélium lui-même au-dessus de 2,2 °K qu'on lui a attribué un nom en propre, *l'hélium II*, pour le distinguer de l'hélium liquide au-dessus de cette température, appelé *hélium I*.

Comme seul l'hélium permet d'étudier les températures voisines du zéro absolu, c'est devenu un élément de première importance en science pure aussi bien qu'appliquée. La quantité que l'on peut en extraire de l'atmosphère est négligeable, et les sources les plus importantes sont les gisements de gaz naturels dans lesquels l'hélium, formé par les désintégrations de l'uranium et du thorium de l'écorce terrestre, peut se retrouver piégé. Le puits le plus riche (au Nouveau-Mexique) produit un gaz qui contient 7,5 % d'hélium.

LA CRYOGÉNIE

Incités par les étranges phénomènes découverts au voisinage du zéro absolu, les physiciens ont naturellement consacré tous leurs efforts à

approcher ce zéro absolu d'aussi près que possible, et à améliorer leur connaissance de ce que l'on appelle maintenant la *cryogénie*. L'évaporation de l'hélium liquide peut, sous certaines conditions, mener à des températures de 0,5 °K. (De telles températures sont mesurées par des méthodes électriques particulières – par exemple par l'intensité du courant créé dans un thermocouple, par la résistance d'un fil de métal conducteur normal, non supraconducteur, par des variations de propriétés magnétiques, ou même par la vitesse du son dans l'hélium. Il n'est guère plus facile de mesurer les très basses températures que de les atteindre.) Des températures sensiblement inférieures ont été atteintes par une méthode initialement proposée en 1925 par le physicien hollandais Petrus Joseph Wilhelm Debye. Un corps *paramagnétique* (c'est-à-dire un corps qui concentre les lignes de champ magnétique) est placé presque au contact de l'hélium liquide, séparé par de l'hélium gazeux, et la température du système est abaissée à environ 1 °K. Le système est ensuite placé dans un champ magnétique. Les molécules du corps paramagnétique s'orientent suivant les lignes de champ et, ce faisant, cèdent de la chaleur. Cette chaleur est emportée par une légère évaporation de l'hélium ambiant. Le champ magnétique est alors supprimé. Les molécules se retrouvent immédiatement orientées au hasard. En passant d'orientations ordonnées à des orientations aléatoires, les molécules absorbent de la chaleur, dont la seule source est l'hélium liquide. La température de l'hélium liquide s'abaisse donc.

On répète ce processus, abaissant ainsi à chaque étape la température, technique qui a été perfectionnée par le chimiste américain William Francis Giauque, ce qui lui a valu le prix Nobel de chimie en 1949. On a de cette façon atteint une température de 0,000 02 °K en 1957.

Le physicien britannique d'origine allemande Heinz London et ses collaborateurs ont imaginé en 1962 une nouvelle méthode pour atteindre des températures encore plus basses. Il existe deux espèces d'hélium, l'*hélium 3* et l'*hélium 4*. Dans les conditions ordinaires elles se mélangent parfaitement ; mais à des températures inférieures à 0,8 °K, elles se séparent, la couche supérieure étant constituée d'hélium 3. Dans la couche inférieure il y a un peu d'hélium 3 avec l'hélium 4, et il est possible de faire aller et venir l'hélium 3 à travers la surface de séparation, ce qui abaisse à chaque fois la température, comme le fait l'échange entre liquide et vapeur dans le cas d'un réfrigérant ordinaire comme le fréon. Les premiers systèmes de refroidissement sur ce principe ont été réalisés en Union Soviétique en 1965.

Le physicien russe Isaak Yakovievitch Pomeranchuk a proposé en 1950 une méthode de refroidissement intense utilisant d'autres propriétés de l'hélium 3, tandis que dès 1934 le physicien britannique d'origine hongroise Nicholas Kurti avait proposé d'utiliser des propriétés magnétiques similaires à celles auxquelles Giauque avait eu recours, mais faisant appel au noyau atomique – le constituant intérieur de l'atome – plutôt qu'aux atomes et molécules entiers.

A l'aide de ces nouvelles techniques on a pu atteindre des températures de 0,000 001 °K. Mais puisqu'ils se trouvent à un millionième de degré du zéro absolu, les physiciens ne pourraient-ils simplement se débarrasser du peu d'entropie qui reste et parvenir finalement à la limite elle-même ?

Non ! Le zéro absolu est inaccessible, comme Nernst l'a montré par ses travaux sur la question (dont la conclusion est parfois appelée le *troisième*

principe de la thermodynamique), récompensés par le prix Nobel. Quel que soit l'abaissement de la température, on ne peut extraire qu'une partie de l'entropie. De façon générale, ôter la moitié de l'entropie d'un système est toujours aussi difficile, quelle qu'en soit la quantité totale. Ainsi, il est aussi difficile de passer de 300 °K (la température ordinaire) à 150 °K (plus bas que n'importe quelle température observée en Antarctique) que de 20 °K à 10 °K. Il est ensuite aussi difficile de descendre de 10 °K à 5 °K, et de 5 °K à 2,5 °K, etc. Étant parvenu à un millionième de degré au-dessus du zéro absolu, atteindre un demi-millionième de degré est aussi coûteux que de passer d'un demi-millionième à un quart de millionième, et ainsi de suite indéfiniment. Quelle que soit la distance parcourue, le zéro absolu reste toujours infiniment éloigné.

Les dernières étapes de cette quête du zéro absolu ont, entre autres, permis d'étudier en détail l'hélium 3, qui est un corps extrêmement rare. L'hélium est déjà en lui-même peu abondant sur la Terre, et lorsqu'on l'isole, seuls 13 atomes sur 10 millions ont des noyaux d'hélium 3, le reste étant constitué d'hélium 4.

L'hélium 3 est un peu plus simple que l'hélium 4 ; sa masse atomique vaut les trois quarts de la masse de la variété la plus commune. Le point de liquéfaction de l'hélium 3 est à 3,2 °K, soit un bon degré au-dessous de celui de l'hélium 4. De plus, on croyait qu'à la différence de l'hélium 4, qui devient superfluide au-dessous de 2,2 °K, l'hélium 3 continuerait à ne montrer aucun signe de superfluidité. Il suffisait de s'obstiner. On a découvert en 1972 que l'hélium 3 se métamorphose au-dessous de 0,0025 °K en une forme liquide superfluide du type hélium II.

LES HAUTES PRESSIONS

La réalisation de hautes pressions a été un des nouveaux horizons scientifiques ouverts par les travaux sur la liquéfaction des gaz. Il semblait qu'en soumettant différents corps (pas seulement des gaz) à des hautes pressions, on pourrait obtenir des informations fondamentales aussi bien sur le comportement de la matière que sur l'intérieur de la Terre. A 10 kilomètres de profondeur, par exemple, la pression est de 1 000 atmosphères ; à 500 kilomètres, de 200 000 atmosphères ; à 3 000 kilomètres, de 1,4 million d'atmosphères ; et au centre de la Terre, à 6 000 kilomètres de profondeur, la pression atteint 3,5 millions d'atmosphères. (Bien sûr, la Terre est plutôt une petite planète. On estime que les pressions au centre de Saturne doivent être supérieures à 50 millions d'atmosphères et dans Jupiter, qui est encore plus grande, 100 millions d'atmosphères.)

Les laboratoires du XIXe siècle ne purent faire mieux que 3 000 atmosphères, pression atteinte par Émile Hilaire Amagat entre 1880 et 1890. Mais en 1905, le physicien américain Percy Williams Bridgman commença à mettre au point de nouvelles méthodes ; elles permirent rapidement d'atteindre des pressions de 20 000 atmosphères, qui firent éclater les minuscules enceintes métalliques qu'il utilisait pour ses expériences. Passant à des matériaux plus résistants, il parvint finalement à produire des pressions d'un demi-million d'atmosphères. Ses travaux sur les hautes pressions ont été récompensés par le prix Nobel de physique en 1946.

Sous ces ultra-hautes pressions, Bridgman réussit à forcer les atomes et les molécules à adopter des dispositions plus compactes qui subsistaient parfois après retour à pression normale. Il a par exemple converti le

phosphore jaune ordinaire, non conducteur de l'électricité, en une forme de phosphore noir et conducteur. Il a induit des changements étonnants, même pour l'eau. La glace ordinaire est moins dense que l'eau liquide. Au moyen de hautes pressions, Bridgman a produit une série de types de glace (la *glace II*, la *glace III*, etc.) non seulement plus denses que le liquide, mais qui restent à l'état de glace à des températures bien supérieures au point de congélation normal de l'eau. La glace VII est encore solide à des températures supérieures au point d'ébullition de l'eau.

Le mot *diamant* évoque le plus fascinant des exploits des hautes pressions. Le diamant bien sûr est du carbone cristallisé, comme le graphite. Lorsqu'un élément peut exister sous deux formes, on appelle celles-ci des *variétés allotropiques*. Le diamant et le graphite en sont l'exemple le plus spectaculaire. L'ozone et l'oxygène ordinaire en sont un autre exemple, ainsi que le phosphore jaune et le phosphore noir (il existe aussi un phosphore rouge), auxquels je faisais allusion dans le paragraphe précédent.

Aspects et propriétés des variétés allotropiques peuvent être totalement différents, et il n'y a pas d'exemple plus frappant que le graphite et le diamant – à l'exception, peut-être, du charbon et du diamant (d'un point de vue chimique, le type de charbon appelé anthracite est une variété désordonnée du graphite).

Il semble à première vue incroyable que le diamant ne soit autre chose que du graphite (ou du charbon) dont les atomes sont disposés différemment, mais le fait fut prouvé il y a longtemps, par Lavoisier et quelques collègues chimistes français en 1772. Ils s'étaient cotisés pour acheter un diamant, qu'ils avaient chauffé à température assez élevée pour le faire brûler. Ils trouvèrent que le gaz résiduel était du dioxyde de carbone. Le chimiste anglais Smithson Tennant montra plus tard, en mesurant la quantité de dioxyde de carbone, que celui-ci n'avait pu être produit que par du carbone pur, ce que devait donc être le diamant, comme le graphite.

Il s'agissait là d'une opération manifestement peu rentable, mais pourquoi ne pas l'inverser ? La densité du diamant est 55 % plus élevée que celle du graphite. Pourquoi ne pas comprimer du graphite et forcer ses atomes à adopter la configuration plus compacte du diamant ?

On a consacré bien des efforts à cette entreprise et, tout comme les alchimistes, plusieurs expérimentateurs ont clamé leur succès. Le cas le plus célèbre est celui du chimiste français Ferdinand Frédéric Henri Moissan. En 1893, après avoir dissous du graphite dans de la fonte en fusion, il annonça avoir trouvé de petits diamants dans la masse de fonte refroidie. La plupart des objets trouvés étaient noirs, minuscules et impurs, mais l'un d'eux était incolore et sa longueur approchait le millimètre. Ces résultats furent généralement acceptés, et on a longtemps estimé que Moissan avait fabriqué des diamants synthétiques. Mais on n'est jamais parvenu à répéter ses résultats.

La recherche des diamants synthétiques a quand même remporté des victoires annexes. En chauffant du graphite dans des conditions qui pourraient, espérait-il, donner du diamant, l'inventeur américain Edward Goodrich Acheson tomba sur le carbure de silicium, auquel il donna le nom commercial de *Carborundum*. Celui-ci s'est avéré plus dur que toute substance connue, sauf le diamant, et c'est, depuis sa découverte, un *abrasif* – c'est-à-dire un agent de meulage ou de polissage – largement utilisé.

L'efficacité d'un abrasif est à la mesure de sa dureté. Un abrasif peut agir sur les corps moins durs que lui-même, et le diamant, le corps le plus dur de tous, est donc le plus efficace. On mesure généralement la dureté des corps sur *l'échelle de Mohs* (ou sur *l'échelle de Brinell* en France), introduite par le minéralogiste allemand Friedrich Mohs en 1818. Cette échelle assigne des degrés aux minéraux, de 1 pour le talc, à 10 pour le diamant. Un corps de degré donné peut rayer tous les minéraux de degré inférieur. Le carborundum a le degré 9 sur l'échelle de Mohs. Mais les divisions ne sont pas égales. Dans l'absolu, la différence de dureté entre le dixième degré (le diamant) et le neuvième (le carborundum) est quatre fois plus grande que la différence entre le neuvième et le premier (le talc).

La raison de ces divers niveaux de dureté n'est guère difficile à trouver. Dans le graphite, les atomes de carbone sont disposés en couches. Dans chaque couche, les atomes sont arrangés selon un réseau hexagonal, comme des carreaux de salle de bains. Chaque atome de carbone est lié de la même façon à trois autres ; et comme l'atome de carbone est petit, les voisins sont très proches et étroitement maintenus. Cet arrangement est difficile à arracher, mais sa finesse le rend facile à casser. Une couche est relativement éloignée de ses voisines supérieure et inférieure en sorte que les couches sont peu liées, et il est facile de faire glisser une couche par rapport à l'autre. Pour cette raison, le graphite n'est pas particulièrement dur, et il est même utilisé comme lubrifiant.

Dans le diamant, par contre, les atomes de carbone sont disposés de façon absolument symétrique dans les trois dimensions. Chaque atome de carbone est lié à *quatre* autres à égale distance, chacun de ceux-ci étant disposé au sommet d'un tétraèdre dont le premier atome occupe le centre. Il s'agit là d'un arrangement très compact qui explique que le diamant soit beaucoup plus dur que le graphite. Et il n'y a aucune direction préférentielle dans laquelle il puisse céder, à moins d'être soumis à une force considérable. D'autres atomes peuvent adopter la *configuration du diamant,* mais de tous ceux-ci ce sont les atomes de carbone qui sont les plus petits et dont l'attraction mutuelle est la plus grande. Dans les conditions qui règnent à la surface de la Terre, le diamant est donc le plus dur de tous les corps.

Dans le carbure de silicium, la moitié des atomes de carbone sont remplacés par des atomes de silicium. Comme les atomes de silicium sont beaucoup plus gros que les atomes de carbone, les voisins ne se serrent plus d'aussi près, et leurs liaisons sont moins fortes. Ainsi, le carbure de silicium n'est pas aussi dur que le diamant (mais il est bien assez dur pour toutes sortes d'utilisations).

Dans les conditions normales, l'arrangement des atomes de carbone du graphite est plus stable que l'arrangement du diamant. Ainsi, le diamant a tendance à se transformer spontanément en graphite. Vous ne courez toutefois aucun risque de vous réveiller un matin pour vous apercevoir que le splendide diamant de votre bague est devenu sans valeur. Même dans leur disposition instable, les atomes de carbone se maintiennent si bien qu'il faudrait des millions d'années pour que la transformation se produise.

C'est cette différence de stabilité qui rend si ardue la transformation du graphite en diamant. Ce n'est qu'en 1930 que les chimistes ont finalement réalisé la pression nécessaire pour convertir le graphite en diamant. On a trouvé que cette conversion exige une pression minimale de 10 000 at-

mosphères, et encore, dans ces conditions, la métamorphose serait d'une lenteur peu pratique. En élevant la température on accélère la conversion, mais il faut une pression plus élevée. A 1 500 °C, une pression minimale de 30 000 atmosphères serait nécessaire. Tout ceci montre que Moissan et ses contemporains, dans les conditions où ils se plaçaient, ne pouvaient pas plus produire de diamants que les alchimistes ne pouvaient fabriquer de l'or. (Quelques indices suggèrent que Moissan fut en fait victime d'un de ses assistants qui, lassé, décida d'en finir avec ses fastidieuses expériences en introduisant un authentique diamant d'origine dans la fonte.)

En s'appuyant sur les travaux de Bridgman pour atteindre des températures et pressions élevées, des chercheurs de la General Electric Company ont finalement réussi en 1955. Il a fallu réaliser des pressions supérieures à 100 000 atmosphères, sous des températures dépassant 2 500 °C. Avec un peu de métal, comme le chrome, on forme un film liquide sur le graphite. C'est au contact de ce film que le graphite se transforme en diamant. En 1962 on a pu atteindre une pression de 200 000 atmosphères à 5 000 °C. Le graphite se transforme alors directement en diamant, sans recourir aux services d'un catalyseur.

Les diamants synthétiques sont trop petits et pas assez purs pour jouer un rôle en joaillerie, mais on les produit maintenant commercialement pour la fabrication des meules et des outils de coupe où ils sont fréquemment employés. On parviendra peut-être bientôt à produire un petit diamant de qualité joaillère.

Un nouveau produit, élaboré par le même type de traitement, peut concurrencer le diamant. Un composé de bore et d'azote (le *nitrure de bore*) a des propriétés tout à fait semblables à celles du graphite (excepté que le nitrure de bore est blanc au lieu d'être noir). Soumis aux mêmes température et pression que celles qui changent le graphite en diamant, le nitrure de bore subit une transformation similaire. Partant d'un arrangement cristallin comme celui du graphite, les atomes de nitrure de bore passent à une disposition analogue à celle du diamant. Sous cette nouvelle forme, on l'appelle *borazon*. Le borazon est environ quatre fois plus dur que le carborundum. Il présente de plus l'immense avantage de mieux résister à la chaleur. A 900 °C, le diamant brûle, mais le borazon résiste.

Le bore a un électron de moins que le carbone, l'azote un électron de plus. Combinés en alternance, ces deux atomes arrivent à reconstituer un arrangement qui ressemble beaucoup à l'arrangement carbone-carbone, avec toutefois un léger écart à la parfaite symétrie du diamant.

Les travaux de Bridgman sur les hautes pressions ne constituent pas, bien sûr, le mot de la fin. Au début des années quatre-vingts, Peter M. Bell de l'Institution Carnegie, utilisant un système qui presse les échantillons entre deux diamants, est parvenu à atteindre des pressions de 1,5 million d'atmosphères, plus de 40 % de la pression qui règne au centre de la Terre. Bell pense que son appareil peut aller jusqu'à 17 millions d'atmosphères avant que les diamants eux-mêmes ne cèdent.

A l'Institut de technologie de Californie (le Caltech), on a produit momentanément, en utilisant des ondes de choc, des pressions encore plus élevées, peut-être de 75 millions d'atmosphères.

Les métaux

La plupart des éléments du tableau périodique sont des métaux. Seulement 20 des 102 éléments peuvent être en fait considérés comme rigoureusement non métalliques. Et pourtant, l'utilisation des métaux n'a fait son apparition que relativement tard dans l'histoire de l'humanité. Une raison en est que dans la nature, à de rares exceptions près, les éléments métalliques sont combinés à d'autres éléments et, par là même, difficiles à reconnaître et à extraire. Les hommes primitifs n'ont d'abord utilisé que des matériaux qui pouvaient être travaillés par des procédés simples comme le taillage, l'écaillage par choc, la coupe et le broyage ; la liste de ces matériaux ne comprenait donc que l'os, la pierre et le bois.

Nos ancêtres ont pu se trouver confrontés aux métaux en découvrant des météorites, des petites pépites d'or, ou du cuivre métallique dans les cendres des feux allumés sur des roches contenant du minerai de cuivre. En tout cas, les gens qui eurent assez de curiosité (et de chance) pour remarquer ces étranges substances nouvelles, et pour chercher des procédés permettant de les traiter, devaient leur trouver de nombreux avantages. Le métal diffère de la roche par l'éclat attrayant qu'un polissage peut lui conférer. On peut le marteler pour en faire des feuilles ou l'étirer en fils. On peut le faire fondre et le couler dans des moules dont il prend la forme. Il est beaucoup plus beau et plus adaptable que la pierre, et c'est le matériau idéal pour la bijouterie. On a probablement façonné des métaux en bijoux bien avant toute autre utilisation.

Étant rares, attirants et inaltérables avec le temps, on attachait de la valeur à ces métaux qui furent troqués tant et si bien qu'ils devinrent des objets d'échange reconnus. Au début, il fallait peser séparément les pièces de métal (or, argent ou cuivre) au cours des transactions commerciales mais, vers le VII[e] siècle av. J.-C., le royaume de Lydie, en Asie Mineure, et les îles de la mer Égée émirent des poids de métal standardisés, frappés d'une estampille officielle du gouvernement. On utilise encore maintenant les pièces de monnaie.

C'est la découverte que certains métaux pouvaient acquérir un tranchant plus effilé que la pierre, et que ce tranchant pouvait résister à un traitement fatal pour une hache de pierre, qui a établi la position définitive de ces métaux. De plus, le métal était robuste. Un choc qui aurait fendu un gourdin en bois ou fait éclater une hache de pierre ne faisait que déformer légèrement un objet métallique de taille comparable. Ces avantages compensaient largement le fait que le métal est plus lourd que la pierre et plus difficile à obtenir.

Le cuivre, utilisé vers l'an 4000 av. J.-C., a été le premier métal obtenu en quantité appréciable. Le cuivre lui-même est trop mou pour en faire des armes ou des blindages efficaces (bien qu'il puisse être utilisé dans un but décoratif), mais on le trouvait souvent allié à un peu d'arsenic ou d'antimoine, d'où il résultait un matériau plus dur que le métal pur. On a dû ensuite découvrir des échantillons de minerai de cuivre qui contenaient de l'étain. La dureté de l'alliage cuivre-étain (le *bronze*) était suffisante pour l'armurerie. Les hommes apprirent vite à ajouter volontairement de l'étain. L'âge de bronze se substitua à l'âge de pierre en Égypte et en Asie occidentale vers l'an 3000 av. J.-C., et dans le Sud-Est de l'Europe vers 2000 av. J.-C. L'*Iliade* et l'*Odyssée* d'Homère commémorent

cette période de l'histoire de l'humanité. Le fer fut découvert à la même époque que le bronze, mais les météorites en restèrent longtemps la seule source connue. Il ne dépassa pas le stade de métal précieux, d'utilisation limitée, tant que restaient à trouver des méthodes d'affinage du minerai de fer pour obtenir celui-ci en quantité illimitée. La difficulté consistait à opérer dans un feu assez chaud avec des méthodes qui permettaient d'ajouter du carbone au fer pour le durcir et obtenir ce que l'on appelle maintenant l'*acier*. La sidérurgie est apparue quelque part en Asie Mineure vers 1400 av. J.-C. et ne s'est développée et propagée que lentement.

Une armée équipée de fer pouvait mettre en déroute une armée équipée de bronze, car les épées de fer coupent le bronze. Les Hittites, en Asie Mineure, furent les premiers à utiliser de façon notable les armes en fer, et ils eurent leur période de domination sur l'Asie occidentale. Les Assyriens prirent ensuite la succession des Hittites. Vers 800 av. J.-C., ils possédaient une armée complètement équipée en fer qui allait dominer le Proche-Orient et l'Égypte pendant deux siècles et demi. A peu près à la même époque, les Doriens importaient l'âge de fer en Europe, en envahissant la Grèce et en battant les Achéens qui avaient commis l'erreur d'en rester à l'âge de bronze.

LE FER ET L'ACIER

Le fer s'obtient essentiellement en chauffant du minerai de fer (généralement un oxyde ferreux) avec du carbone. Les atomes de carbone emportent l'oxygène de l'oxyde de fer, laissant un lingot de fer pur. Au début, les températures atteintes ne permettaient pas de fondre le fer, et on obtenait un métal robuste que l'on pouvait façonner à la forme désirée par martelage – c'était le *fer forgé*. La sidérurgie, la métallurgie du fer, apparut à grande échelle au Moyen Age. On utilisait des fours spéciaux et des plus hautes températures qui fondaient le fer. Le fer fondu pouvait être coulé dans des moules pour former des pièces de fonderie ; on l'a donc appelé *fonte*. Ce matériau était beaucoup moins coûteux que le fer forgé et beaucoup plus dur, mais il était fragile et ne se prêtait pas au martelage. La demande croissante en fer sous ses deux formes contribua à déboiser l'Angleterre par exemple, car on utilisait le bois comme combustible dans les fourneaux à fondre le fer. Mais en 1780, le sidérurgiste anglais Abraham Darby montra que le *coke* (le charbon distillé) convenait aussi bien, si ce n'est mieux, que le *charbon de bois* (le bois carbonisé). La demande imposée aux forêts s'allégea et le charbon devint, pour plus d'un siècle, la source d'énergie dominante.

Ce n'est qu'à la fin du XVIIIe siècle que les chimistes réalisèrent, grâce au physicien français René Antoine Ferchault de Réaumur, que c'était finalement la teneur en carbone qui dictait au fer sa dureté et sa résistance. Pour maximiser ces qualités, la teneur en carbone doit être comprise entre 0,2 et 1,5 % ; l'acier qui en résulte est généralement plus dur, plus robuste et plus résistant que la fonte ou le fer forgé. Mais jusqu'au milieu du XIXe siècle, on ne pouvait fabriquer de l'acier de bonne qualité que par le procédé compliqué qui consistait à ajouter soigneusement la quantité convenable de carbone au fer forgé (lui-même relativement coûteux). L'acier restait donc un métal de luxe, utilisé seulement lorsqu'il n'existait aucun substitut, comme pour les épées et les ressorts.

C'est un ingénieur britannique, nommé Henry Bessemer, qui introduisit l'âge de l'acier. Initialement intéressé surtout par les canons et les projectiles, Bessemer avait inventé un système de canon rayé qui devait améliorer la portée et la précision des canons. L'empereur Napoléon III était intéressé par le projet et proposa de subventionner d'autres expériences. Mais un artilleur français anéantit cette idée en faisant remarquer que l'explosion propulsive à laquelle songeait Bessemer ferait éclater les canons en fonte alors en service. Déçu, Bessemer se mit à rechercher comment fabriquer un fer plus résistant. Ne connaissant rien à la métallurgie, il pouvait aborder la question avec un esprit neuf. C'était sa teneur en carbone qui rendait la fonte fragile. Le problème consistait donc à abaisser le taux de carbone.

Pourquoi ne pas brûler ce carbone en fondant le fer et en faisant passer un courant d'air à travers celui-ci ? Au premier abord, c'était une idée ridicule. L'air soufflé ne refroidirait-il pas le métal fondu, pour finalement le solidifier ? Bessemer essaya quand même et s'aperçut que c'était en fait le contraire qui se produisait. A mesure que l'air brûlait le carbone, la combustion produisait de la chaleur et la température du minerai augmentait au lieu de baisser. Le carbone brûlait tranquillement. En agissant convenablement, on avait un procédé relativement bon marché pour produire l'acier en grande quantité !

En 1856 Bessemer présenta au public son *haut fourneau.* Les fabricants d'acier adoptèrent avec enthousiasme ce système puis, mécontents, l'abandonnèrent quand ils s'aperçurent que l'acier produit était de qualité inférieure. Bessemer réalisa que le minerai de fer utilisé dans l'industrie contenait du phosphore (contrairement à ses propres échantillons). Bessemer eut beau expliquer aux sidérurgistes que le phosphore était la cause de leur infortune, ils refusèrent de se laisser attraper une seconde fois. Bessemer dut emprunter de l'argent et monter sa propre usine à Sheffield. Important du minerai de fer sans phosphore de Suède, il parvint rapidement à fabriquer de l'acier à un prix inférieur à celui des autres producteurs.

En 1875, le sidérurgiste britannique Sidney Gilchrist Thomas découvrit qu'en doublant l'intérieur du haut fourneau avec de la chaux et de la magnésie, on pouvait facilement se débarrasser du phosphore contenu dans le fer en fusion. Depuis, on peut utiliser presque n'importe quel minerai de fer pour produire de l'acier. Entre-temps, l'inventeur britannique d'origine allemande Karl Wilhelm von Siemens avait mis au point, en 1868, le *four à sole ouverte,* dans lequel des lingots de fer étaient chauffés avec le minerai, ce qui réglait également le problème du phosphore.

L'âge de l'acier avait maintenant commencé. Ce n'est pas simple rhétorique. Sans l'acier, les gratte-ciel, les ponts suspendus, les grands navires, les chemins de fer et bien d'autres réalisations modernes seraient presque inimaginables et, en dépit de l'usage croissant d'autres métaux, l'acier reste le matériau préférentiel dans quantités d'applications quotidiennes, des carrosseries automobiles à la coutellerie.

(C'est bien sûr une erreur de croire qu'une seule découverte peut apporter une révolution majeure dans l'évolution de l'humanité. Un tel changement est toujours le résultat de nombreux développements interconnectés. Avec tout l'acier du monde, par exemple, on ne pourrait conférer une valeur pratique aux gratte-ciel, sans l'existence de l'ascenceur, un système trop

souvent considéré comme implicitement acquis. L'inventeur américain Elisha Graves Otis a breveté en 1861 un ascenseur hydraulique ; et en 1889 la compagnie qu'il avait fondée a installé les premiers ascenceurs électriques dans un immeuble commercial de New York.)

L'acier, abondant et bon marché, permettait de procéder expérimentalement à l'addition d'autres métaux (pour obtenir des *aciers alliés,* afin de voir si on pouvait encore l'améliorer). C'est le métallurgiste britannique Robert Abbott Hadfield qui a inauguré cette voie. Il a découvert en 1882 que l'addition de 13 % de manganèse à l'acier fournissait un alliage plus dur, utilisable en construction mécanique lorsqu'interviennent des efforts brutaux, comme dans le concassage des pierres. En 1900, on a trouvé qu'un alliage d'acier, de tungstène et de chrome conservait sa dureté à haute température, même porté au rouge ; cet alliage, appelé *acier rapide,* a été providentiel pour l'industrie de l'outillage. Aujourd'hui, il y a quantité d'autres aciers alliés, faisant appel à des métaux comme le molybdène, le nickel, le cobalt et le vanadium, adaptés à des fonctions particulières.

Le gros problème avec l'acier est sa sensibilité à la corrosion, un processus qui ramène le fer à son état primitif originel de minerai. Une défense consiste à protéger le métal par une couche de peinture ou d'un métal plus résistant à la corrosion, tel que le nickel, le chrome, le cadmium ou l'étain. Former un alliage qui ne se corrode pas est une méthode plus efficace. Le métallurgiste britannique Harry Brearley a découvert accidentellement cet alliage en 1913. Il était à la recherche d'aciers alliés pour confectionner des canons de fusil. Parmi ses échantillons d'essai abandonnés il y avait un alliage nickel-chrome. Des mois plus tard, il s'aperçut que dans son tas de ferrailles, ces morceaux en particulier étaient toujours aussi brillants alors que tout le reste avait rouillé. Ainsi était né *l'acier inoxydable.* Il est trop mou et trop coûteux pour être utilisable à grande échelle en construction, mais convient admirablement en coutellerie et pour tous les petits appareils où l'inaltérabilité est plus importante que la dureté.

Depuis que l'on dépense quelques milliards de dollars par an à travers le monde en efforts infructueux pour empêcher le fer et l'acier de se corroder, la recherche d'un produit antirouille universel se poursuit sans arrêt. On a fait récemment l'intéressante découverte que les *pertechnétates* (des composés du technétium) protègent le fer contre la rouille. Il est bien entendu probable que cet élément rare, produit seulement en laboratoire, ne sera jamais assez abondant pour être utilisable à l'échelle industrielle, mais il n'en constitue pas moins un inestimable outil de recherche. Sa radioactivité permet aux chimistes de le suivre à la trace et d'observer en particulier ce qui lui arrive à la surface du fer.

Son fort ferromagnétisme est l'une des propriétés du fer les plus utiles. Le fer lui-même est un exemple d'*aimant* faible. On le magnétise facilement sous l'influence d'un champ magnétique, c'est-à-dire que ses domaines magnétiques (voir le chapitre 5) sont facilement alignables. Il se démagnétise aussi aisément lorsqu'on supprime le champ, et les domaines retombent dans des orientations au hasard. Cette possibilité d'annuler le magnétisme peut être pratique, par exemple dans les *électro-aimants,* où on magnétise facilement le noyau de fer en faisant passer le courant, mais où il doit pouvoir se démagnétiser tout aussi facilement lorsqu'on interrompt ce courant.

Depuis la Deuxième Guerre mondiale, on a mis au point une nouvelle classe de matériaux aimantables. Ce sont les *ferrites*, par exemple la ferrite au nickel ($NiFe_2O_4$) et la ferrite au manganèse ($MnFe_2O_4$), utilisées dans les ordinateurs pour leur aptitude à se démagnétiser extrêmement rapidement et facilement.

Les *aimants permanents* – dont les domaines sont difficiles à orienter et qui, une fois orientés, sont tout aussi difficiles à remettre dans le désordre –, une fois magnétisés, conservent cette propriété pendant longtemps. Différents aciers alliés en constituent l'exemple le plus ordinaire, bien qu'on soit parvenu à réaliser des aimants permanents particulièrement forts avec des alliages contenant peu, ou même pas du tout de fer. Le meilleur exemple connu est l'*alnico*, découvert en 1931, dont une variété est faite d'aluminium, de nickel et de cobalt (comme son nom l'indique), plus un peu de cuivre.

Durant les années cinquante, on a mis au point des techniques faisant appel à du fer en poudre en guise d'aimant, les particules étant si petites qu'elles forment individuellement des domaines. Ces particules peuvent être orientées dans un plastique liquide qui, une fois solidifié, maintient les domaines dans leur orientation initiale. Ces *aimants plastiques* sont très malléables, mais on peut également en réaliser des variétés dures.

LES NOUVEAUX MÉTAUX

Ces dernières décennies ont vu l'apparition de nouveaux métaux particulièrement utiles. L'exemple le plus frappant en est l'aluminium. L'aluminium est le métal le plus commun ; il est 60 % plus abondant que le fer. Mais il est très difficile à extraire de son minerai. En 1825, Hans Christian Oersted (qui avait découvert le lien entre électricité et magnétisme) avait séparé un peu d'aluminium sous une forme encore impure. A partir de là, bien des chimistes tentèrent sans succès de purifier ce métal, jusqu'à ce que le chimiste français Henri Etienne Sainte-Claire Deville mette finalement au point, en 1854, une méthode pour obtenir de l'aluminium pur en quantité appréciable. L'aluminium est chimiquement tellement actif que Sainte-Claire Deville dut recourir au sodium métallique (encore plus actif) pour briser l'étreinte de l'aluminium avec les atomes voisins. Pendant un certain temps le prix de l'aluminium en fit pratiquement un métal précieux. Napoléon III appréciait les couteaux en aluminium et fit confectionner un hochet en aluminium pour son fils héritier ; aux Etats-Unis, en hommage de la nation à George Washington, le monument de celui-ci fut surmonté d'une plaque d'aluminium massif en 1885.

En 1886, Charles Martin Hall, un jeune étudiant en chimie, fut si impressionné par le propos de son professeur, qui prédisait la fortune à quiconque découvrirait un procédé bon marché pour produire l'aluminium, qu'il décida de tenter sa chance. Dans son laboratoire personnel, aménagé dans une cabane en planches, il se proposa d'appliquer la découverte antérieure de Humphry Davy, à savoir qu'un courant électrique traversant un métal en fusion permet de séparer les ions métalliques par dépôt sur la cathode. Cherchant une substance qui dissoudrait l'aluminium, il tomba sur la *cryolite*, un minéral que l'on ne trouve en quantité appréciable qu'au Groenland. (On dispose maintenant de cryolite synthéti-

que.) Hall prépara un mélange d'oxyde d'aluminium et de cryolite qu'il fit fondre, et à travers lequel il fit passer un courant électrique. Comme prévu, l'aluminium pur vint se déposer sur la cathode. Hall se précipita chez son professeur avec ses premiers lingots d'aluminium. (Ces lingots sont encore précieusement conservés par l'Aluminum Company of America.)

Il se trouve qu'un jeune chimiste français nommé Paul Louis Toussaint Héroult, qui avait juste l'âge de Hall (vingt-deux ans), découvrit le même procédé la même année. (Pour compléter la coïncidence, Hall et Héroult sont morts tous les deux en 1914.)

Bien que le *procédé Hall-Héroult* ait fait de l'aluminium un métal peu coûteux, celui-ci n'a jamais été aussi bon marché que l'acier, car le minerai d'aluminium utilisable est moins abondant que le minerai de fer équivalent, et parce que l'électricité (indispensable pour l'aluminium) est plus coûteuse que le charbon (utilisé pour l'acier). L'aluminium présente néanmoins deux avantages importants sur l'acier. Il est d'abord léger ; il ne pèse qu'un tiers du poids de l'acier. Deuxièmement, la corrosion de l'aluminium se limite à un film superficiel, fin et transparent, qui protège les couches les plus profondes de la corrosion sans affecter l'apparence du métal.

L'aluminium pur est relativement mou, mais l'alliage peut y remédier. Le métallurgiste allemand Alfred Wilm a préparé, en 1906, un alliage dur en ajoutant un peu de cuivre et une quantité encore plus faible de magnésium. Il a vendu les droits d'exploitation de son brevet à la société métallurgique Durener, en Allemagne, qui a donné à cet alliage le nom de duralumin.

Les ingénieurs ont vite réalisé l'intérêt d'un métal léger, mais résistant, pour l'aviation. Après l'introduction par les Allemands du duralumin dans les dirigeables, et la découverte de sa composition par l'analyse de l'alliage trouvé dans les restes d'un dirigeable écrasé, l'usage de ce nouveau métal s'est répandu à travers le monde. Le duralumin n'étant pas tout à fait aussi résistant à la corrosion que l'aluminium, les métallurgistes le recouvrent d'une couche fine d'aluminium, pour former un produit appelé alclad.

Il existe maintenant des alliages d'aluminium qui, à poids égal, ont une résistance plus élevée que certains aciers. L'aluminium tend à prendre la place de l'acier chaque fois que légèreté et résistance à la corrosion sont plus importantes que la simple résistance mécanique. Comme chacun sait, l'aluminium est devenu un métal presque universel, utilisé dans les avions, les fusées, le matériel roulant, les automobiles, les huisseries et les cloisons de maisons, les peintures, les ustensiles de cuisine, les feuilles d'emballage, etc.

Et nous en venons au *magnésium*, un métal encore plus léger que l'aluminium. Comme on pouvait s'y attendre, son principal domaine d'application est l'aviation ; dès 1910 l'Allemagne utilisait à cette fin des alliages magnésium-zinc. Après la Première Guerre mondiale on a de plus en plus utilisé des alliages magnésium-aluminium.

La quantité disponible de magnésium n'étant que le quart de celle de l'aluminium, et le magnésium se trouvant chimiquement plus actif, il est plus difficile à extraire de ses minerais. Heureusement, les océans en constituent un réservoir très riche. A la différence de l'aluminium ou du fer, le magnésium se trouve en grande quantité dans l'eau de mer. Les corps dissous dans l'Océan représentent 3,5 % de sa masse. Dans ces

matériaux en solution, il y a 3,7 % d'ions magnésium. Au total les océans contiennent environ deux mille billions (2 000 000 000 000 000) de tonnes de magnésium, soit toute la consommation envisageable pour l'éternité.

Le problème est d'extraire ce magnésium. La méthode choisie consiste à pomper l'eau de mer dans de grands réservoirs pour y ajouter de l'oxyde de calcium (également fourni par la mer, dans les coquilles d'huîtres). L'oxyde de calcium réagit avec l'eau et l'ion magnésium pour former l'hydroxyde de magnésium insoluble, qui précipite donc de la solution. Un traitement à l'acide chlorhydrique convertit l'hydroxyde de magnésium en chlorure de magnésium, et le magnésium métal est enfin séparé du chlore au moyen d'un courant électrique.

En janvier 1941, la Dow Chemical Company a produit les premiers lingots de magnésium à partir de l'eau de mer, point de départ du décuplement de la production de magnésium durant les années de guerre.

A vrai dire, tout élément que l'on peut économiquement extraire de l'eau de mer peut être considéré comme effectivement inépuisable car, après utilisation, il finit par retourner à la mer. On estime qu'avec une extraction annuelle de 100 millions de tonnes de magnésium à partir de l'eau de mer, la teneur en magnésium de l'Océan tomberait, après un million d'années d'exploitation, de 0,13 % à 0,12 %.

Si l'acier a été le métal miracle du milieu du XIXe siècle, l'aluminium celui du début du XXe siècle et le magnésium celui du milieu, quel sera le prochain métal miracle ? Les possibilités sont limitées. Il n'y a que sept métaux réellement communs dans l'écorce terrestre. A côté du fer, de l'aluminium et du magnésium, il y a le sodium, le potassium, le calcium et le titane. Chimiquement, le sodium, le potassium et le calcium sont beaucoup trop actifs pour être utilisés en construction. (En particulier, ils réagissent violemment avec l'eau.) Il ne reste que le titane, dont l'abondance est environ un huitième de celle du fer.

Le titane offre une extraordinaire combinaison de qualités. Son poids ne dépasse guère la moitié de celui de l'acier ; à poids égal il est plus résistant que l'aluminium ou l'acier ; il résiste à la corrosion et peut supporter de hautes températures. Pour toutes ces raisons on utilise maintenant le titane dans les avions, les navires et les missiles guidés, chaque fois qu'il peut être fait bon usage (!) de ses propriétés.

Pourquoi la reconnaissance des mérites du titane a-t-elle été si lente ? La cause en est tout à fait similaire à ce qui s'est produit pour l'aluminium et le magnésium : le titane réagit trop facilement avec les autres corps et, lorsqu'il est impur – combiné à l'oxygène ou à l'azote –, c'est un métal peu avenant, fragile et apparemment inutile. Sa résistance et ses autres qualités ne se manifestent que lorsqu'il est isolé sous forme réellement pure (dans le vide ou en atmosphère inerte). Les efforts des métallurgistes ont à tel point réussi que le prix du titane est passé de 6 000 dollars le kilo en 1947 à 4 dollars en 1969.

Mais il n'est pas nécessaire de chercher de nouveaux métaux miracles. On peut rendre les vieux métaux (ainsi que quelques corps non métalliques) beaucoup plus « miraculeux » qu'ils ne le sont actuellement.

Un poème de Wendel Holmes, « Le chef-d'œuvre du diacre », raconte l'histoire d'un « char-à-joual » soigneusement construit pour n'avoir aucun point faible. Le cabriolet à un cheval finit par se pulvériser tout d'un coup. Mais il a duré cent ans.

La structure atomique des solides cristallins, qu'il s'agisse de métaux ou non, fait penser à cette situation. Les cristaux d'un métal sont criblés de fissures et crevasses microscopiques. Lorsque le cristal est soumis à une pression, une fracture prend naissance en l'un de ces points faibles et se propage à travers le cristal. Un cristal que l'on pourrait fabriquer, comme la merveilleuse voiture du diacre, sans point faible, jouirait d'une résistance supérieure.

Il existe en fait de tels cristaux sans points faibles qui se forment en fibres minuscules appelées *whiskers* (« moustaches de chat ») à la surface des cristaux. On a mesuré la résistance de fibres de carbone qui allait jusqu'à 2 000 kilos par millimètre carré, soit de 15 à 70 fois la limite de rupture en traction des aciers. Si l'on parvenait à mettre au point des méthodes pour fabriquer des métaux sans défauts en quantité industrielle, on disposerait de matériaux de résistance étonnante. Les chercheurs soviétiques, par exemple, sont parvenus en 1968 à produire un minuscule cristal de tungstène sans défauts qui pouvait supporter 2 500 kilos par millimètre carré, à comparer avec les 350 kilos par millimètre carré du meilleur acier. Et même si ces corps sans défauts ne pouvaient être obtenus en grosse quantité d'une seule pièce, l'addition de fibres sans défauts aux métaux ordinaires renforcerait ceux-ci.

Mais, aussi récemment qu'en 1968, on a encore trouvé une nouvelle méthode intéressante pour combiner les métaux. Les deux méthodes consacrées par l'histoire étaient l'*alliage,* lorsque plusieurs métaux sont fondus ensemble et forment un mélange plus ou moins homogène, et la *métallisation,* dans laquelle un métal est solidement appliqué contre un autre (généralement on applique une couche fine d'un métal précieux à la surface d'une masse de métal bon marché ; ainsi la surface peut être aussi élégante et inaltérable que l'or, par exemple, tandis que l'objet est lui-même presque aussi bon marché que le cuivre).

Le métallurgiste américain Newell C. Cook et ses associés ont tenté d'appliquer une couche de silicium sur une surface de platine, en plongeant le platine dans un bain de fluorure alcalin en fusion. La métallisation espérée ne s'est pas produite. Apparemment, le fluorure en fusion détruisait le film très fin d'oxygène lié, ordinairement présent même sur les métaux les plus résistants, et exposait à nu la surface du platine aux atomes de silicium. Au lieu de s'attacher à la surface par l'intermédiaire des atomes d'oxygène, les atomes de silicium pénétraient carrément dans la surface. Une fine couche externe de platine devenait ainsi un alliage.

Cook a suivi cette nouvelle direction et a trouvé que l'on peut combiner bien des corps de cette façon pour former un « placage » d'alliage sur métal pur (ou sur un autre alliage). Cook a appelé ce procédé la *métallidation* (« metalliding ») et rapidement démontré son utilité. Ainsi, le cuivre auquel on ajoute de 2 à 3 % de béryllium, sous forme d'alliage ordinaire, devient extraordinairement résistant. Mais on peut obtenir le même résultat en « béryllidant » le cuivre, ce qui exige bien moins de béryllium, relativement rare. De même, l'acier métallidé au bore (« boridé ») est durci. L'addition de silicium, de cobalt ou de titane confère également des propriétés utiles.

En d'autres termes, si on ne les trouve pas dans la nature, les métaux miraculeux peuvent être créés par l'ingéniosité des hommes.

Chapitre 7

Les particules

L'atome nucléé

Comme je l'ai mentionné dans le chapitre précédent, on savait en 1900 que l'atome n'était pas une particule simple, indivisible, mais qu'il contenait au moins un type de particule subatomique, l'électron, identifié par J. J. Thomson. Thomson suggéra que les électrons étaient entassés comme des raisins secs dans l'ensemble de l'atome proprement dit, chargé positivement.

L'IDENTIFICATION DES PARTICULES

Mais on s'aperçut rapidement qu'il y avait d'autres particules dans l'atome. Lorsque Becquerel découvrit la radioactivité, il identifia certains des rayonnements émis par les corps radioactifs comme étant des électrons ; mais il y avait d'autres types de rayonnements. Les époux Curie en France et Ernest Rutherford en Angleterre trouvèrent que l'un de ces rayonnements avait moins de pouvoir de pénétration que les électrons. Rutherford baptisa ce nouveau rayonnement le *rayonnement alpha* et donna à l'émission d'électrons le nom de *rayonnement bêta*. Les électrons libres qui constituent ce dernier sont individuellement des *particules bêta*. On trouva que le rayonnement alpha était lui aussi formé de particules ; celles-ci furent appelées *particules alpha*. *Alpha* et *bêta* sont les deux premières lettres de l'alphabet grec.

Pendant ce temps, le chimiste français Paul Ulrich Villard découvrait une troisième forme d'émission radioactive, évidemment baptisée *rayonne-*

ment gamma, d'après la troisième lettre de l'alphabet grec. On identifia rapidement les rayons gamma ; ils ressemblaient aux rayons X, mais avec des longueurs d'onde plus courtes.

Rutherford trouva, par ses expériences, qu'un champ magnétique déviait les particules alpha, mais beaucoup moins que les particules bêta. De plus, les déviations étaient opposées ; la particule alpha avait donc une charge positive, opposée à la charge négative de l'électron. On pouvait calculer, d'après la valeur de la déviation, que la masse de la particule alpha devait être au moins double de celle de l'ion hydrogène, la plus petite charge positive connue. En effet, la déviation dépendait de la masse de la particule et de sa charge ; si la charge positive de la particule alpha était égale à celle de l'ion hydrogène, sa masse devait être le double de la masse de celui-ci ; si sa charge était double, sa masse devait valoir quatre fois celle de l'ion hydrogène, etc. (figure 7.1).

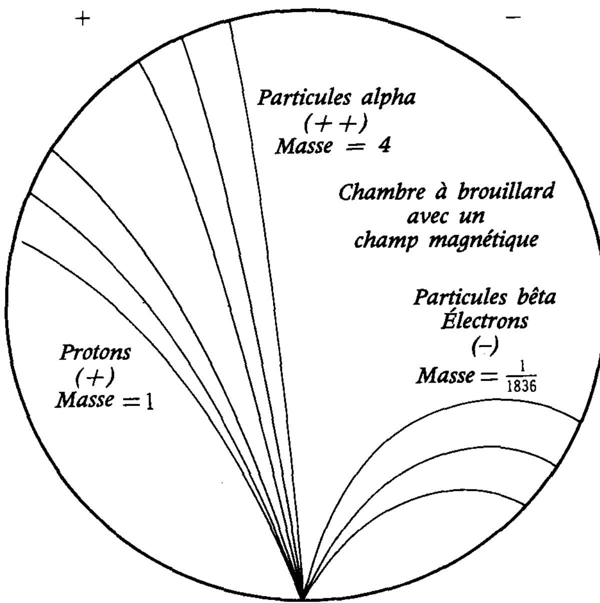


Figure 7.1. Déviations des particules par un champ magnétique.

Rutherford régla la question en 1909 en isolant les particules alpha. Il disposait un corps radioactif dans une ampoule de verre mince, elle-même placée dans un tube vide en verre épais. Les particules alpha pouvaient traverser la mince paroi de verre intérieure, mais pas la paroi extérieure épaisse. Elles rebondissaient, pour ainsi dire, sur la paroi extérieure et, ce faisant, perdaient de l'énergie au point de devenir incapables de pénétrer à nouveau la paroi mince. Les particules étaient donc piégées entre les deux parois. Rutherford excitait ensuite les particules alpha avec une décharge électrique, ce qui les faisait émettre de la lumière. Ce rayonnement avait les mêmes raies spectrales que celui de l'hélium. (On sait maintenant que ce sont les particules alpha produites par les corps

radioactifs dans le sol qui sont à l'origine de l'hélium des gisements de gaz naturel.) A supposer que la particule alpha soit de l'hélium, alors sa masse devait être quatre fois plus grande que celle de l'hydrogène et sa charge positive s'élever à deux unités, en prenant comme unité la charge de l'ion hydrogène.

Rutherford identifia ensuite une autre particule positive dans l'atome. Celle-ci avait en fait été détectée, sans être identifiée, bien des années auparavant. En utilisant un tube à rayons cathodiques avec une cathode perforée, le physicien allemand Eugen Goldstein avait découvert, en 1886, un nouveau rayonnement qui sortait par les trous de la cathode dans le sens opposé aux rayons cathodiques eux-mêmes. Il les avait baptisés *rayons canaux*. Ce rayonnement servit dès 1902, lorsqu'on mit en évidence l'effet Doppler-Fizeau (voir le chapitre 2) dans toute source lumineuse terrestre. Le physicien allemand Johannes Stark, après avoir placé un spectroscope de façon que les rayons se dirigent vers celui-ci, montra leur décalage vers le violet. Ses travaux furent récompensés par le prix Nobel de physique en 1919.

Comme les rayons canaux se déplaçaient en sens inverse des rayons cathodiques négativement chargés, Thomson proposa d'appeler ce rayonnement les *rayons positifs*. Il s'avéra que les particules des rayons positifs traversaient facilement la matière. On en déduisit qu'elles étaient donc beaucoup plus petites que les ions ou atomes ordinaires. La valeur de leur déviation par un champ magnétique indiquait que la plus petite de ces particules avait la même charge et la même masse que l'ion hydrogène, en supposant que cet ion était porteur de la plus petite quantité possible de charge positive. D'où il découlait que la particule des rayons positifs était la particule positive fondamentale, le pendant de l'électron. Rutherford l'appela *proton* (d'après le mot grec pour « premier »).

Le proton et l'électron portent effectivement les mêmes charges électriques (quoique opposées), bien que le proton soit 1 836 fois plus lourd que l'électron. Il semblait donc probable qu'un atome soit composé de protons et d'électrons compensant mutuellement leurs charges. On avait également tout lieu de supposer que les protons se trouvaient au cœur de l'atome car, contrairement aux électrons, il est difficile de les en extraire. La grande question était alors la suivante : quel type de structure ces particules de l'atome forment-elles ?

LE NOYAU ATOMIQUE

C'est Rutherford qui tomba sur un début de réponse. De 1906 à 1908, il ne cessa de soumettre au bombardement de particules alpha des feuilles métalliques minces (de l'or ou du platine), pour en sonder les atomes. la plupart des projectiles traversaient sans être déviés (comme des balles de fusil traversent le feuillage d'un arbre). Mais il y avait des exceptions : en examinant la plaque photographique qui servait de cible derrière la feuille métallique, Rutherford s'aperçut qu'il y avait une dispersion inattendue des impacts autour de la tache centrale, et même que les particules rebondissaient vers l'arrière ! Tout se passait comme si certaines des balles n'avaient pu traverser le feuillage sans encombre, mais avaient ricoché sur quelque chose de plus consistant.

Rutherford estima qu'il avait tapé sur une sorte de noyau dur qui n'occupait qu'une toute petite partie du volume de l'atome. La plus grande

partie du volume de l'atome devait, semblait-il, être occupée par les électrons. Quand une particule alpha traversait la feuille métallique, elle ne rencontrait le plus souvent que des électrons et écartait cette brume de particules légères sans être elle-même déviée. Mais de temps en temps il arrivait qu'une particule alpha frappe le noyau plus dense d'un atome et se trouve alors déviée. Que ce fait soit relativement rare montrait que les noyaux d'atomes devaient être réellement très petits, car un projectile passant à travers la feuille métallique rencontrait des milliers d'atomes.

On pouvait, en bonne logique, supposer que le noyau dur était fait de protons. Rutherford représenta les protons d'un atome bien serrés en un minuscule *noyau atomique* au centre de l'atome. (On a montré depuis que le diamètre de ce noyau est à peine supérieur à 1/100 000 de celui de l'atome total.)

Tel est donc le modèle de base de l'atome : un noyau chargé positivement qui ne prend que très peu de place, mais qui contient presque toute la masse de l'atome, environné d'un brouillard d'électrons représentant à peu près tout le volume de l'atome, mais ne contenant pratiquement rien de sa masse. Pour cet extraordinaire travail de précurseur sur la nature fondamentale de la matière, Rutherford reçut le prix Nobel de chimie en 1908.

Il devenait alors possible de décrire des atomes donnés et leur comportement en termes mieux définis. Par exemple, l'atome d'hydrogène ne possède qu'un seul électron. Si on ôte celui-ci, le proton restant s'attache immédiatement à une molécule voisine. Mais lorsque ce noyau d'hydrogène nu ne trouve aucun électron à partager de cette façon, il se comporte comme un proton – c'est-à-dire une particule subatomique ; sous cette forme il peut pénétrer la matière et, s'il a l'énergie suffisante, réagir avec d'autres noyaux.

L'hélium, affublé de deux électrons, n'en laisse pas partir un facilement. Comme je l'ai dit dans le précédent chapitre, ses deux électrons constituent une couche complète, et l'atome est donc inerte. Mais, si l'hélium est dépouillé de ses deux électrons, il devient une particule alpha, c'est-à-dire une particule subatomique porteuse de deux unités de charge positive.

L'atome du troisième élément, le lithium, comporte trois électrons. Dépouillé d'un ou deux de ses électrons, c'est un ion. Débarrassé de ses trois électrons, il devient à son tour un noyau nu, porteur de trois unités de charge positive.

L'atome total étant habituellement neutre, le nombre d'unités de charge positive du noyau d'un atome doit être rigoureusement égal au nombre d'électrons qu'il contient normalement. Les numéros atomiques des éléments sont en fait fondés sur leur quantité de charges positives, plutôt que négatives, car en formant un ion on peut facilement faire varier le nombre d'électrons d'un atome, tandis qu'on ne peut en modifier le nombre de protons que difficilement.

Ce schéma de construction des atomes était à peine élaboré qu'une nouvelle question surgissait. Le nombre d'unités de charge positive d'un noyau ne correspondait pas du tout à sa masse, excepté dans le cas de l'atome d'hydrogène. Le noyau d'hélium, par exemple, a deux charges positives, mais on savait que sa masse était de quatre fois celle du noyau d'hydrogène. Et la situation empirait à mesure que l'on progressait dans le tableau des éléments, jusqu'à l'uranium dont la masse est égale à celle de 238 protons et à laquelle ne correspond qu'une charge de 92.

Comment un noyau qui contenait quatre protons (comme on le supposait du noyau d'hélium) pouvait-il n'avoir que deux unités de charge positive ? La première idée, et la plus simple, était que deux unités de sa charge étaient compensées par la présence dans le noyau de particules chargées négativement de masse négligeable. C'était naturellement l'électron qui se présentait à l'esprit. Le problème pouvait être réglé en admettant que le noyau d'hélium était fait de quatre protons et deux électrons compensateurs, ce qui laissait une charge positive nette de deux – et ainsi de suite jusqu'à l'uranium, dont le noyau devait être formé de 238 protons et 146 électrons, soit 92 unités de charge positive. Ce point de vue était renforcé par le fait que l'on savait que des noyaux radioactifs émettent des électrons (particules bêta).

Ce modèle prévalut durant plus d'une décennie, jusqu'à ce qu'une meilleure réponse survienne, de façon détournée, de recherches dans un domaine différent. Mais entre-temps, de sérieuses objections à cette hypothèse s'étaient fait jour. En particulier, si le noyau était essentiellement constitué de protons, les électrons légers ne contribuant pratiquement pas à sa masse, comment se faisait-il que les masses relatives des divers noyaux ne se réduisaient pas à des nombres entiers ? D'après les mesures de poids atomique, la masse de l'atome de chlore, par exemple, valait 35,5 fois celle du noyau d'hydrogène. Contenait-il donc 35,5 protons ? Aucun chercheur (à l'époque, ni maintenant) ne pouvait accepter cette idée d'une moitié de proton.

En fait, c'est la réponse à cette dernière question qui fut trouvée, avant même la solution au problème principal. C'est une histoire intéressante en elle-même.

Les isotopes

DES BRIQUES UNIFORMES

Dès 1816, le physicien anglais William Prout avait suggéré que tous les atomes étaient construits à partir de l'atome d'hydrogène. Au fur et à mesure que le temps passait et que l'on déterminait les poids atomiques, la théorie de Prout était tombée dans l'oubli car de nombreux éléments se révélaient avoir des poids atomiques fractionnaires (en prenant l'oxygène comme référence avec la valeur 16). Le chlore, on l'a dit, a un poids atomique de 35,453. Mais il y a d'autres exemples : l'antimoine (121,75), le baryum (137,34), le bore (10,811), le cadmium (112,40).

Au tournant du siècle, on fit une série d'observations étonnantes qui devaient finalement conduire à l'explication. L'Anglais William Crookes (celui du tube de Crookes) avait séparé de l'uranium une petite quantité d'un corps qui s'avérait beaucoup plus radioactif que l'uranium lui-même. Il émit l'idée que l'uranium n'était pas radioactif du tout, que seule l'était cette impureté qu'il appela *uranium X*. Henri Becquerel, d'un autre côté, avait découvert que l'uranium purifié, faiblement radioactif, voyait sa radioactivité augmenter quelque peu au cours du temps. Après l'avoir laissé tranquille quelque temps, on pouvait toujours en extraire à nouveau de l'uranium X. En d'autres termes, l'uranium était converti, par sa propre radioactivité, en uranium X encore plus actif.

De façon similaire, Rutherford avait ensuite séparé le *thorium X*, fortement radioactif, du thorium, et trouvé que le thorium, lui aussi, continuait à produire plus de thorium X. On savait déjà que le plus célèbre des éléments radioactifs, le radium, donnait un gaz radioactif, le radon. Aussi Rutherford et son assistant Frederick Soddy conclurent-ils que les atomes radioactifs se transformaient généralement en atomes radioactifs d'autres sortes, en même temps qu'ils émettaient leurs particules.

Les chimistes se mirent à la recherche de ces transformations et se retrouvèrent avec un assortiment de nouveaux corps, leur donnant des noms comme *radium A, radium B, mésothorium I, mésothorium II et actinium C*. Tous ceux-ci étaient groupés en trois séries, selon leur ascendance atomique. Une série venait de l'éclatement de l'uranium, une autre du thorium et une troisième de l'actinium (l'actinium s'avéra plus tard avoir lui-même un prédécesseur, le *protoactinium*). On identifia en tout une quarantaine de membres de ces séries, chacun se distinguant par son mode de rayonnement particulier. Mais le produit final des trois séries était le même : chaque chaîne de corps finissait par se terminer au même élément stable, le plomb.

Evidemment, ces quarante substances ne pouvaient chacune être un élément séparé ; entre l'uranium (92) et le plomb (82), il n'y avait que dix places dans la classification périodique, toutes occupées, sauf deux, par des éléments connus. Les chimistes constatèrent en fait que, bien que différentes par leur radioactivité, certaines de ces substances étaient identiques en ce qui concernait leurs propriétés chimiques. Par exemple, dès 1907, les chimistes américains Herbert Newby McCoy et William Horace Ross montrèrent que le *radiothorium*, l'un des produits de la désintégration du thorium, manifestait exactement le même comportement chimique que le thorium. Le *radium D* se comportait chimiquement exactement comme le plomb ; en fait, il était souvent appelé *radioplomb*. Tout ceci indiquait que les substances en question étaient en réalité des variétés d'un même élément : le radiothorium, une forme de thorium ; le radioplomb, un membre de la famille des plombs ; et ainsi de suite.

En 1913, Soddy exprima clairement cette idée et la poursuivit. Il montra que lorsqu'un atome émet une particule alpha, il se change en un élément situé deux cases plus bas dans le tableau périodique des éléments ; lorsqu'il émet une particule bêta, il est changé en l'élément de la case au-dessus. Selon ce schéma, le radiothorium tombait bien à la place du thorium dans le tableau, ainsi que les corps appelés *uranium* X_1 et *uranium Y* : ils étaient tous trois des variétés de l'élément 90. De même, le radium D, le radium B, le thorium B et l'actinium B partageaient tous la case du plomb et étaient de variétés de l'élément 82.

Aux membres d'une famille de substances partageant la même position dans le tableau périodique, Soddy a donné le nom d' *isotopes* (du mot grec signifiant « même position »). Il fut lauréat du prix Nobel de chimie 1921.

Le modèle du noyau à base de protons et d'électrons – qui finalement se révéla faux – s'adaptait parfaitement à la théorie des isotopes de Soddy. L'enlèvement d'une particule alpha d'un noyau devait réduire la charge positive de ce noyau de deux unités, exactement ce qu'il fallait pour le faire reculer de deux cases dans le tableau périodique. D'un autre côté, l'éjection d'un électron (une particule bêta) d'un noyau devait laisser un proton non compensé de plus et donc augmenter la charge positive du noyau d'une unité. Ceci avait pour effet d'augmenter le numéro atomique

d'une unité, de sorte que l'élément passait à la case immédiatement au-dessus dans le tableau périodique.

Comment se fait-il que lorsque le thorium devient du radiothorium, après non pas une, mais trois désintégrations, le produit soit encore du thorium ? Dans ce processus, l'atome de thorium perd une particule alpha, puis une particule bêta, et enfin une deuxième particule bêta. Si l'on accepte l'idée du proton comme élément de construction de base, l'atome de thorium a perdu quatre électrons (dont deux qui devraient être contenus dans la particule alpha) et quatre protons. (La réalité diffère de cette représentation, mais d'une façon qui n'affecte pas le résultat.) Nous avons supposé que le noyau a commencé avec 232 protons et 142 électrons. Ayant perdu quatre protons et quatre électrons, il est ramené à 228 protons et 138 électrons. Dans tous les cas, le nombre de protons non compensés (232-142, ou 228-138) est de 90. Ceci nous laisse donc avec le numéro atomique 90, le même qu'au départ. Ainsi le radiothorium, comme le thorium, possède 90 électrons planétaires qui orbitent autour du noyau. Les propriétés chimiques d'un atome étant déterminées par le nombre de ses électrons planétaires, le thorium et le radiothorium ont chimiquement le même comportement, quelle que soit la différence de leurs poids atomiques (232 contre 228).

On identifie les isotopes d'un élément par leur poids atomique, ou leur *nombre de masse*. Ainsi le thorium ordinaire est appelé *thorium 232*, tandis que le radiothorium est le *thorium 228*. De même, les isotopes radioactifs du plomb sont connus sous les dénominations de *plomb 210* (radium D), *plomb 214* (radium B), *plomb 212* (thorium B) et *plomb 211* (actinium B).

La notion d'isotope s'applique aux éléments stables aussi bien qu'aux éléments radioactifs. Par exemple, les trois séries radioactives que j'ai mentionnées se terminent par trois sortes de plomb différentes. La série de l'uranium s'achève avec le plomb 206, la série du thorium avec le plomb 208 et la série de l'actinium avec le plomb 207. Chacun de ceux-ci est un isotope stable « ordinaire » du plomb, mais ces trois plombs diffèrent par leur poids atomique.

La preuve de l'existence d'isotopes stables a été apportée par un appareil inventé par Francis William Aston, un assistant de J.J. Thomson. Il s'agissait d'un système très sensible pour séparer les isotopes grâce aux différences de déviations de leurs ions dans un champ magnétique. Aston baptisa cet appareil *spectrographe de masse*. Avec un prototype de cet instrument, Thomson montra en 1919 que l'atome de néon existait sous deux variétés : l'une avec un nombre de masse de 20, l'autre avec un nombre de masse de 22. Le néon 20 était l'isotope ordinaire, le néon 22 se trouvait associé au néon 20 dans la proportion de un atome sur dix. (On découvrit plus tard un troisième isotope, le néon 21, présent dans le néon de l'atmosphère, dans la proportion de un atome sur quatre cents.)

La raison des valeurs fractionnaires pour les poids atomiques des éléments devenait enfin claire. Le poids atomique de 20,183 pour le néon représentait la masse résultante des trois isotopes différents qui interviennent dans l'élément tel qu'on le trouve dans la nature. Chaque atome individuel avait un nombre de masse entier, mais le nombre de masse moyen, le poids atomique, était fractionnaire.

Aston continua en montrant que plusieurs éléments stables ordinaires étaient en fait des mélanges d'isotopes. Il trouva que le chlore, avec un poids atomique fractionnaire de 35,453, était constitué de chlore 35 et de

chlore 37, avec une *abondance relative* de 3 pour 1. Aston a été récompensé par le prix Nobel de chimie en 1922.

Dans sa conférence, lors de la réception du prix, Aston prévoyait explicitement la possibilité d'utiliser l'énergie contenue dans le noyau atomique, anticipant ainsi les centrales et les armes nucléaires (voir le chapitre 10). Le physicien américain d'origine canadienne, Arthur Jeffrey Dempster, fit en 1935 un pas important dans cette direction en utilisant l'appareil d'Aston. Il montra que, bien que 993 atomes d'uranium sur 1 000 soient de l'uranium 238, les sept atomes restants sont de l'uranium 235. Cette découverte était lourde d'une signification que l'on n'allait pas tarder à comprendre.

Ainsi, après un siècle de fausses pistes, l'idée de Prout était enfin reconnue. Les éléments *sont* construits avec des composants identiques qui, s'ils ne sont pas des atomes d'hydrogène, ont du moins une masse égale à celle de l'atome d'hydrogène. La raison pour laquelle les éléments ne manifestent pas cette construction modulaire dans leurs poids tient au fait que l'on a affaire à des mélanges d'isotopes dont chacun contient un nombre différent de composants constitutifs. En réalité, même l'oxygène, dont le poids atomique de 16 était utilisé comme référence pour mesurer les poids relatifs des éléments, n'est pas un cas parfaitement pur. Pour 10 000 atomes d'oxygène 16 ordinaire, on trouve vingt atomes d'un isotope dont la masse est de 18 unités et quatre atomes dont le nombre de masse vaut 17.

Très peu d'éléments, en fait, ont un isotope unique. (L'expression est impropre : dire d'un élément qu'il a un seul isotope équivaut à dire d'une femme qu'elle a donné naissance à « un seul jumeau ».) Dans cette catégorie d'éléments, on trouve le béryllium, dont tous les atomes ont un nombre de masse qui vaut 9, le fluor, constitué uniquement de fluor 19 ; l'aluminium, uniquement d'aluminium 27 ; et un certain nombre d'autres. Un noyau de structure donnée est maintenant parfois appelé un *nuclide*, suivant une proposition du chimiste américain Truman Paul Kohman faite en 1947. On peut donc parfaitement dire qu'un élément tel que l'aluminium est constitué d'un seul nuclide.

SUR LES TRACES DES PARTICULES

Depuis que Rutherford a identifié la première particule nucléaire (la particule alpha), les physiciens n'ont cessé de fouiller dans le noyau, essayant de changer un atome en un autre, ou s'efforçant de le briser pour voir de quoi il est fait. Ils ne disposaient au début que de la particule alpha, dont Rutherford fit un excellent usage.

L'une des expériences fructueuses que Rutherford et ses assistants réalisèrent consistait à envoyer des particules alpha sur un écran couvert de sulfure de zinc. Chaque impact produisait une minuscule scintillation (un effet qui avait été découvert par Crookes en 1903) de sorte que l'on pouvait observer à l'œil nu et compter les arrivées des particules individuelles. Poursuivant cette technique, les expérimentateurs interposèrent un disque de métal qui devait empêcher les particules alpha d'atteindre le sulfure de zinc et donc faire cesser toute scintillation. Mais en introduisant de l'hydrogène dans l'appareil, on voyait encore scintiller l'écran en dépit de la présence du disque de métal. De plus, ces nouvelles

scintillations avaient un aspect différent de celles produites par les particules alpha. Comme le disque de métal arrêtait les particules alpha, il devait y avoir d'autres radiations qui passaient à travers pour atteindre l'écran. On se persuada que ce rayonnement devait être constitué de protons rapides. En d'autres termes, les particules alpha devaient taper de temps en temps en plein dans le noyau d'un atome d'hydrogène (qui, souvenons-nous-en, consiste en un proton) et le projeter en avant, comme le choc d'une boule de billard peut en expédier une autre vers l'avant. Les protons frappés, relativement légers, devaient être projetés à grande vitesse ; ils pouvaient donc pénétrer le disque métallique et atteindre l'écran de sulfure de zinc.

Cette détection des particules individuelles par scintillation est un exemple de *compteur à scintillations*. Lors de leurs expériences, Rutherford et ses assistants devaient d'abord rester un quart d'heure dans l'obscurité pour accoutumer leurs yeux, puis s'astreindre à des comptages pénibles. Les compteurs à scintillations modernes ne font plus appel à l'œil et au cerveau humain. Les scintillations sont converties en impulsions électriques qui sont ensuite comptées par des circuits électroniques. Il n'y a plus qu'à lire le résultat final sur un affichage convenable. Lorsque les scintillations sont nombreuses, on peut procéder de façon plus pratique en utilisant des circuits électroniques qui n'enregistrent qu'une impulsion sur deux ou sur quatre (ou même plus). Ces *échelles* (qui réduisent le comptage, pour ainsi dire) ont été mises au point par le physicien anglais Charles Eryl Wynn-Williams en 1931. Depuis la Deuxième Guerre mondiale, on a avantageusement substitué des corps organiques au sulfure de zinc.

Les expériences originales de Rutherford prirent un tour totalement inattendu lorsque, substituant de l'azote à de l'hydrogène en guise de cible soumise au bombardement des particules alpha, il constata que l'écran de sulfure de zinc présentait encore des scintillations exactement semblables à celles produites par les protons. La seule conclusion qui s'offrait à Rutherford était que le bombardement éjectait des protons des noyaux d'azote.

Pour essayer de déterminer exactement ce qui arrivait, Rutherford utilisa alors une *chambre à brouillard de Wilson*, un système inventé en 1895 par le physicien écossais Charles Thomson Rees Wilson. Un récipient de verre, muni d'un piston, est rempli d'air saturé d'humidité. Lorsqu'on tire le piston, l'air subit une détente soudaine et donc se refroidit. A température plus basse, il est sursaturé d'humidité. Dans de telles conditions, toute particule chargée va constituer un germe de condensation de la vapeur d'eau. Une particule qui fonce à travers la chambre ionise des atomes et laisse donc un sillage marqué par une ligne brumeuse de gouttelettes.

La nature de cette trace peut en dire beaucoup sur la particule. Une particule bêta, légère, laisse un discret chemin sinueux, car un choc entre cette particule et un électron a les mêmes effets sur les deux protagonistes. Une particule alpha, beaucoup plus massive, laisse une trace rectiligne épaisse. Si elle capture deux électrons pour devenir un atome d'hélium neutre, sa trace se termine. Outre la taille et l'allure de sa trace, il y a d'autres moyens d'identifier une particule dans une chambre de Wilson. Sa réponse à un champ magnétique que l'on peut appliquer nous dit si sa charge est positive ou négative, et la courbure de sa trajectoire nous renseigne sur sa masse et son énergie. Les physiciens sont maintenant

tellement habitués aux photographies de toutes sortes de traces qu'ils peuvent les interpréter à livre ouvert. Wilson fut un des lauréats du prix Nobel de physique 1927 pour sa contribution à la mise au point de la chambre à brouillard.

On a modifié de plusieurs façons la chambre de Wilson depuis son invention, et plusieurs instruments apparentés ont été imaginés. La chambre à brouillard originale était inutilisable après la détente, tant qu'elle n'était pas réinitialisée. Alexander Langsdorf a inventé, en 1939 aux Etats-Unis, une *chambre à brouillard à diffusion* dans laquelle une vapeur chaude d'alcool diffuse dans une région plus froide de sorte qu'il y ait toujours une région sursaturée et que l'on puisse continuellement observer les traces.

Il y a eu ensuite la *chambre à bulles*, un système dont le principe est similaire. Dans ce détecteur on utilise un liquide surchauffé sous pression plutôt qu'un gaz sursaturé. Le trajet de la particule chargée est révélé par une ligne de bulles de vapeur dans le liquide, au lieu de gouttelettes liquides dans la vapeur. On raconte que son inventeur, le physicien américain Donald Arthur Glaser, en eut l'idée en 1953 en contemplant un verre de bière. Si c'est le cas, il s'agit là d'un verre de bière auquel le monde des physiciens doit beaucoup ; Glaser aussi d'ailleurs, puisque le prix Nobel de physique lui a été décerné en 1960 pour l'invention de la chambre à bulles.

La première chambre à bulles n'avait que quelques centimètres de diamètre. En une dizaine d'années, on en est venu à utiliser des chambres à bulles de deux mètres de longueur. Comme les chambres à brouillard à diffusion, les chambres à bulles sont toujours prêtes. Mais comme il y a beaucoup plus d'atomes dans un volume donné de liquide que dans un gaz, la production d'ions est plus importante dans les chambres à bulles, qui se trouvent donc particulièrement bien adaptées à l'étude des particules rapides à courte durée de vie. Dix ans après leur invention, les chambres à bulles produisaient des centaines de milliers de clichés par semaine. On a découvert dans les années soixante des particules à durée de vie ultra-courte qui seraient passées inaperçues sans la chambre à bulles.

L'hydrogène liquide convient parfaitement au remplissage des chambres à bulles car le noyau d'hydrogène, constitué d'un seul proton, est assez simple pour n'introduire qu'un minimum de complications. On a construit en 1973, à Wheaton, Illinois, une chambre à bulles de cinq mètres de diamètre contenant 25 000 litres d'hydrogène liquide. Certaines chambres à bulles contiennent de l'hélium liquide.

En dépit d'une sensibilité aux particules à courte durée de vie supérieure à celle de la chambre de Wilson, la chambre à bulles présente certains inconvénients. Contrairement à la chambre à brouillard, la chambre à bulles ne peut être déclenchée par des événements choisis. Elle ne fait qu'enregistrer globalement, et c'est parmi d'innombrables traces qu'il faut chercher celles qui ont une signification. Il fallait donc trouver une méthode de détection des traces qui combinât sélectivité de la chambre à brouillard et sensibilité de la chambre à bulles.

Ce besoin fut finalement satisfait par la *chambre à étincelles*, où les particules incidentes ionisent du néon et créent des courants électriques dans ce gaz compartimenté par de nombreuses plaques métalliques. Les courants se manifestent sous forme de lignes d'étincelles visibles qui marquent le passage des particules, et le système peut être réglé pour ne réagir qu'aux particules étudiées. La première chambre à étincelles utilisable pratiquement a été construite en 1959 par les physiciens japonais

Saburo Fukui et Shotaro Miyamoto. Des physiciens soviétiques ont amélioré le système en 1963, en augmentant sa sensibilité et sa souplesse. On produit de courts jets de lumière qui, finalement, arrivent à former une ligne pratiquement continue (plutôt que les étincelles séparées de la chambre à étincelles). Ce système modifié est appelé *chambre à streamers* (les « streamers » servent à désigner toutes sortes de bannières, flammes ou oriflammes ainsi que les rayons de soleil bas sur l'horizon). Il permet de détecter des événements qui se produisent dans la chambre et les particules qui en jaillissent dans toutes les directions, difficultés sur lesquelles achoppait la chambre à étincelles originale.

LA TRANSMUTATION DES ÉLÉMENTS

Mais, laissant de côté l'étude perfectionnée des particules subatomiques en vol, il nous faut revenir un demi-siècle en arrière pour voir ce qui arrivait lorsque Rutherford bombardait des noyaux d'azote avec des particules alpha dans une des premières chambres de Wilson. La particule alpha laissait une trace qui s'achevait soudainement par une fourche – à coup sûr, une collision avec un noyau d'azote. Une branche de la fourche, relativement fine, représentait un proton éjecté. L'autre branche, une trace courte et épaisse, représentait ce qui restait du noyau d'azote reculant sous l'effet du choc. Mais il n'y avait aucun signe de la particule alpha elle-même. Elle semblait avoir été absorbée par le noyau d'azote, hypothèse vérifiée plus tard par le physicien britannique Patrick Maynard Stuart Blackett, dont on dit qu'il prit plus de 20 000 clichés pour arriver à trouver huit collisions (assurément un exemple de patience, de confiance et d'obstination surhumaines). Pour cela, et d'autres travaux dans le domaine de la physique nucléaire, Blackett a reçu le prix Nobel de physique en 1948.

Du coup, on comprenait ce qui arrivait au noyau d'azote. En absorbant la particule alpha, son nombre de masse et sa charge positive passaient respectivement de 14 à 18 et de 7 à 9. Cette combinaison perdait immédiatement un proton, le nombre de masse retombait à 17 et la charge positive à 8. Mais l'élément qui a 8 unités de charge positive est l'oxygène, et le nombre de masse 17 désigne l'isotope oxygène 17. En d'autres termes, Rutherford avait, en 1919, transmuté l'azote en oxygène. C'était la première transmutation réalisée de main d'homme. Le rêve des alchimistes était accompli, mais d'une manière qu'ils n'auraient pu prévoir, et encore moins mettre en œuvre avec leurs techniques primitives.

En tant que projectiles, les particules alpha des sources radioactives avaient leur limite : elles n'étaient pas tout à fait assez énergétiques pour pénétrer les noyaux des éléments lourds, dont la charge positive élevée repousse fortement les particules chargées positivement. Mais la forteresse nucléaire était battue en brèche et des attaques plus énergiques allaient survenir.

Les nouvelles particules

Cette allusion aux attaques subies par le noyau nous ramène à la question de la constitution même de ce noyau. La théorie de la structure nucléaire à base de protons et d'électrons, si elle expliquait parfaitement les isotopes,

butait sur quelques autres faits. Les particules subatomiques ont en général une propriété appelée *spin*, quelque chose comme la rotation propre des objets astronomiques sur leur axe. Il se trouve qu'avec les unités utilisées pour mesurer ce spin, les protons et les électrons ont des valeurs de spin de plus ou moins un demi. Ainsi, un nombre pair d'électrons ou de protons (ou des deux) confinés dans un noyau doit doter celui-ci d'un spin dont les valeurs peuvent être 0 ou un nombre entier : + 1, – 1, + 2, – 2, etc. Le spin total d'un noyau constitué d'un nombre d'électrons et de protons impair doit être demi-entier, comme + 1/2, – 1/2, + 3/2, – 3/2, + 5/2, – 5/2, etc. Essayez d'ajouter un nombre pair de moitiés positives ou négatives (ou un mélange) ; faites-en autant avec un nombre impair ; vous verrez qu'il en est ainsi et pas autrement.

Le noyau d'azote a une charge électrique de + 7 et une masse de 14 unités. D'après la théorie proton-électron, ce noyau devait contenir 14 protons pour rendre compte de la masse observée, et 7 électrons pour neutraliser la moitié des charges et laisser un bilan de + 7. Le nombre total de particules dans ce noyau aurait dû être de 21, et le spin résultant du noyau d'azote aurait dû être demi-entier. Mais ce n'était pas le cas, le noyau d'azote avait un spin entier.

Ce genre de désaccord se produisait également dans le cas d'autres noyaux ; la théorie proton-électron semblait donc ne pouvoir convenir. Mais tant qu'ils ne connaissaient que ces particules subatomiques, les physiciens restèrent complètement désarmés pour trouver une nouvelle théorie.

LE NEUTRON

En 1930, deux physiciens allemands, Walther Bothe et Herbert Becker, révélèrent que le noyau émettait un rayonnement nouveau et mystérieux, dont le pouvoir de pénétration était surprenant. Ils avaient produit ce rayonnement par bombardement d'atomes de béryllium avec des particules alpha. Un an auparavant, Bothe avait mis au point une méthode utilisant simultanément deux compteurs ou plus, des *compteurs de coïncidences*. Ces compteurs permettaient d'identifier des événements nucléaires qui se déroulaient en un millionième de seconde. Pour cela, ainsi que pour d'autres travaux, Bothe fut l'un des lauréats du prix Nobel de physique 1954.

Deux ans plus tard, la découverte de Bothe et Becker fut suivie par celle des physiciens français Irène et Frédéric Joliot-Curie. (Irène était la fille de Pierre et Marie Curie, et Joliot avait ajouté le nom de sa femme au sien en l'épousant.) Utilisant la nouvelle radiation du béryllium pour bombarder de la paraffine, une substance cireuse composée d'hydrogène et de carbone, ils constatèrent que cette radiation éjectait des protons de la paraffine.

Le physicien anglais James Chadwick suggéra immédiatement que ce rayonnement était formé de particules. Pour déterminer les caractéristiques de ces particules, il les envoya sur des atomes de bore ; mesurant l'augmentation de masse du noyau formé, il put alors calculer que la masse de la particule incidente sur le bore était de l'ordre de la masse du proton. Et pourtant on ne parvenait pas à détecter la particule proprement dite dans une chambre de Wilson. Chadwick décida donc que l'explication

devait résider dans la neutralité électrique de la particule (une particule non chargée ne produit aucune ionisation et ne condense donc aucune gouttelette d'eau).

Ainsi Chadwick conclut à l'existence d'une particule entièrement nouvelle, une particule qui avait une masse voisine de celle du proton, mais sans aucune charge ou, en d'autres termes, électriquement neutre. La possibilité d'une telle particule avait déjà été envisagée, et on avait même proposé un nom, le *neutron*. Chadwick accepta ce nom et fut récompensé par le prix Nobel de physique 1935 pour la découverte du neutron.

La nouvelle particule éclaircit immédiatement certains doutes que les physiciens théoriciens avaient pu avoir à propos du modèle proton-électron pour le noyau. Le physicien théoricien allemand Werner Heisenberg proclama que l'idée d'un noyau constitué de protons et de neutrons, plutôt que de protons et d'électrons, fournissait une représentation beaucoup plus satisfaisante. On pouvait ainsi considérer le noyau d'azote comme constitué de sept protons et de sept neutrons. Le nombre de masse était alors de 14, et la charge totale (le nombre atomique) de + 7. De plus, le nombre total de particules dans le noyau s'élevait à 14, un nombre pair, au lieu de 21, un nombre impair, dans l'ancienne théorie. Comme le neutron, à l'instar du proton, a un spin de + 1/2 ou − 1/2, un nombre pair de neutrons et protons donnait au noyau d'azote un spin entier, ce qui correspondait aux observations. Tous les noyaux dont les spins étaient inexplicables dans le cadre de la théorie proton-électron avaient finalement des spins qui pouvaient s'expliquer par la théorie proton-neutron. Cette théorie fut finalement acceptée et elle l'est restée. Il n'y a pas d'électrons dans le noyau.

En outre, le nouveau modèle rendait compte du tableau périodique tout aussi bien que l'ancien. Le noyau d'hélium consiste en deux protons et deux neutrons, ce qui explique sa masse de quatre unités et sa charge électrique de deux. L'idée prenait en compte les isotopes de façon très simple. Le noyau de chlore 35, par exemple, a 17 protons et 18 neutrons ; le chlore 37, 17 protons et 20 neutrons. Ils ont donc la même charge électrique, et l'excédent de poids de l'isotope plus lourd réside entièrement dans ses deux neutrons supplémentaires. De même, les trois isotopes de l'oxygène ne diffèrent que par leur nombre de neutrons : l'oxygène 16 a 8 protons et 8 neutrons ; l'oxygène 17, 8 protons et 9 neutrons ; l'oxygène 18, 8 protons et 10 neutrons (figure 7.2).

Pour résumer, disons que chaque élément peut être défini simplement par le nombre de protons de son noyau, qui est équivalent au numéro atomique. Mais tous les éléments, sauf l'hydrogène, ont aussi des neutrons dans leur noyau, et le nombre de masse d'un nuclide est la somme de ses nombres de protons et de neutrons. Ainsi le neutron est associé au proton en tant qu'élément constitutif de base de la matière. On trouve maintenant commode de les désigner tous les deux par le terme générique de *nucléon*, utilisé pour la première fois en 1944 par le physicien danois Christian Moller.

De ce progrès dans la connaissance de la structure nucléaire a résulté une classification supplémentaire des nuclides. Les nuclides qui ont le même nombre de protons sont, comme je l'ai dit, des isotopes. De même, des nuclides qui ont le même nombre de neutrons (comme, par exemple, l'hydrogène 2 et l'hélium 3, dont les noyaux contiennent chacun un

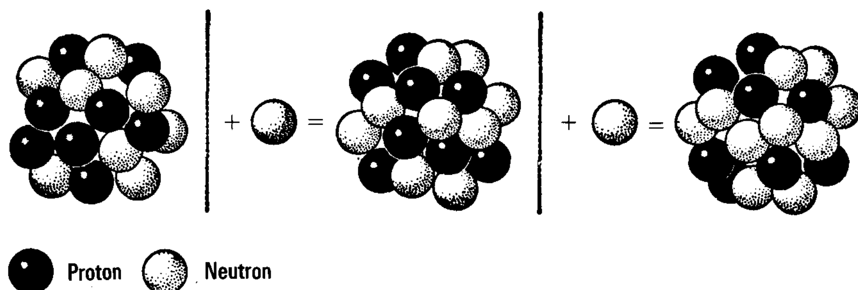


Figure 7.2. La construction des noyaux d'oxygène 16, d'oxygène 17 et d'oxygène 18. Chacun de ces noyaux contient huit protons et respectivement huit, neuf et dix neutrons.

neutron) sont des *isotones*. Les nuclides qui ont le même nombre de nucléons, et donc le même nombre de masse, comme le calcium 40 et l'argon 40, sont des *isobares*.

Il est clair que la théorie proton-neutron de la structure nucléaire était fondamentalement incapable d'expliquer le fait que les noyaux radioactifs puissent émettre des particules bêta (des électrons) : d'où venaient ces électrons, s'il n'y en avait aucun dans le noyau ? On a éclairci cette énigme, comme je vais l'expliquer brièvement.

LE POSITRON

La découverte du neutron déçut les physiciens sur un point très important. Ils étaient parvenus à imaginer un Univers construit avec seulement deux particules fondamentales, le proton et l'électron. Il fallait maintenant en ajouter une troisième. Pour les scientifiques, toute entorse à la simplicité est haïssable.

Et le pire était que cela ne faisait que commencer. Cette exception tourna rapidement à la déroute totale. D'autres particules allaient surgir.

Depuis longtemps, les physiciens étudiaient les mystérieux *rayons cosmiques* de l'espace, découverts en 1911 par le physicien autrichien Victor Francis Hess au cours d'ascensions aérostatiques dans la haute atmosphère.

Ce rayonnement avait été détecté avec un instrument d'une simplicité encourageante pour ceux qui ont parfois l'impression que la science contemporaine ne peut progresser qu'en faisant appel à des appareils d'une incroyable complexité. Cet instrument était un *électroscope*, qui consistait en deux morceaux de feuille d'or fine, fixés à une tige métallique dans une boîte métallique munie de fenêtres. (L'ancêtre de ce système avait été construit longtemps avant, en 1706, par le physicien anglais Francis Hauksbee.)

Lorsque la tige métallique est chargée d'électricité statique, les morceaux de feuille d'or s'écartent. Ils pourraient, en théorie, rester à jamais séparés, mais les ions de l'atmosphère environnante emportent lentement la charge, de sorte que les deux feuilles se rapprochent progressivement. Un rayonnement d'énergie élevée – comme des rayons X, des rayons gamma,

ou un faisceau de particules chargées – produit les ions qui permettent une telle fuite de la charge. Même lorsque l'électroscope est soigneusement isolé, il y a toujours une légère fuite qui dénote la présence d'un rayonnement très pénétrant sans lien direct avec la radioactivité. C'était ce rayonnement pénétrant que Hess détecta et dont l'intensité augmentait à mesure qu'il s'élevait dans l'atmosphère. Hess fut un des lauréats du prix Nobel de physique 1936 pour cette découverte.

Le physicien américain Robert Andrews Millikan, après avoir accumulé une quantité de renseignements sur ce rayonnement (auquel il donna le nom de *rayons cosmiques*), en arriva à la conclusion qu'il devait s'agir d'une forme de rayonnement électromagnétique. Son pouvoir de pénétration était tel qu'un mètre de plomb en laissait encore passer une partie. Pour Millikan, ceci indiquait que ce rayonnement était apparenté aux rayons gamma, mais avec une longueur d'onde encore plus courte.

D'autres, comme le physicien américain Arthur Holly Compton, maintenaient que les rayons cosmiques étaient des particules. Pourtant il y avait un moyen de trancher la question. S'il s'agissait de particules chargées, elles devaient être déviées par le champ magnétique terrestre lorsqu'elles arrivaient, depuis l'espace lointain, au voisinage de la Terre. Compton analysa les mesures de rayonnement cosmique à différentes latitudes et trouva que celui-ci était effectivement courbé par le champ magnétique : le rayonnement était plus faible près de l'équateur magnétique et plus intense près des pôles, où les lignes de champ magnétique plongent vers la Terre.

Lorsqu'elles pénètrent dans l'atmosphère, les particules cosmiques *primaires* ont une énergie incroyablement élevée. La plupart sont des protons, mais on y trouve des noyaux d'éléments plus lourds. D'une façon générale, plus le noyau est lourd, plus il est rare dans les particules cosmiques. Des noyaux aussi complexes que celui du fer furent bientôt détectés, et on a même détecté en 1968 des noyaux d'uranium. Ces noyaux d'uranium ne représentent qu'une particule sur dix millions. On trouve aussi quelques électrons de très haute énergie.

Lorsque ces particules primaires frappent les atomes et les molécules de l'air, elles en fracassent les noyaux et produisent toutes sortes de particules *secondaires*. C'est ce rayonnement secondaire (encore très énergétique) que l'on détecte près du sol, mais des ballons ont enregistré le rayonnement primaire dans la haute atmosphère.

La nouvelle particule suivante – après le neutron – fut découverte à la suite de recherches sur les rayons cosmiques. Un physicien théoricien avait en fait prévu cette découverte. Selon le raisonnement de Paul Adrien Maurice Dirac, suite à une analyse mathématique des propriétés des particules subatomiques, à chaque particule devait correspondre une *antiparticule*. (Les scientifiques aiment que la nature soit non seulement simple, mais symétrique.) Il devait donc exister un *antiélectron,* exactement semblable à l'électron, excepté sa charge positive au lieu de négative, et un *antiproton* avec une charge négative plutôt que positive.

Lorsque Dirac proposa sa théorie, en 1930, elle ne créa pas grand remue-ménage dans le monde scientifique. Mais, comme prévu, l'anti-électron se révéla deux ans plus tard. Le physicien américain Carl David Anderson travaillait avec Millikan sur le problème de la nature corpusculaire ou électromagnétique des rayons cosmiques. A l'époque, la plupart des physiciens acceptaient la preuve apportée par Compton qu'il

s'agissait de particules chargées, mais Millikan était terriblement mauvais perdant, et il ne voulait pas s'avouer battu. Anderson avait pour projet de déterminer si les rayons cosmiques étaient courbés par un fort champ magnétique dans une chambre de Wilson. Pour ralentir suffisamment ces rayons afin que leur courbure, si elle existait, soit mesurable, Anderson disposa dans la chambre un écran de plomb d'un demi-centimètre d'épaisseur. Il trouva que le rayonnement cosmique qui traversait la chambre, après être passé à travers le plomb, laissait une trace courbée. Au cours de leur passage à travers le plomb, les rayons cosmiques de haute énergie expulsaient des particules des atomes de plomb. Une de ces particules laissait une trace exactement semblable à celle d'un électron. Mais elle était courbée dans le mauvais sens ! Même masse, mais charge opposée. C'était lui... l'antiélectron de Dirac. Anderson baptisa sa découverte le *positron.* Il constitue un exemple de rayonnement secondaire produit par les rayons cosmiques ; mais on a découvert en 1963 que les positrons figuraient aussi dans le rayonnement primaire.

Livré à lui-même, le positron est aussi stable que l'électron (pourquoi pas puisque, excepté sa charge électrique, il est identique à l'électron ?) et pourrait exister indéfiniment. Mais il n'est pas abandonné à lui-même puisqu'il est mis au monde dans un univers rempli d'électrons. Dans sa course, il se trouve presque immédiatement (disons en un millionième de seconde) au voisinage d'un électron.

Il peut y avoir, pendant un moment, une association électron-positron, une configuration où les deux particules tournent l'une autour de l'autre. Le physicien américain Arthur Edward Ruark a proposé, en 1945, d'appeler *positronium* ce système à deux particules, et en 1951 le physicien américain d'origine autrichienne Martin Deutsch est parvenu à mettre en évidence le positronium par son spectre de rayons gamma caractéristique.

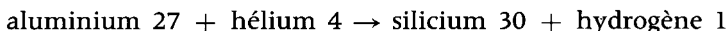
Mais, même si un système comme le positronium se forme, son existence ne dure au plus qu'1/10 000 000 de seconde. La danse de l'électron et du positron s'achève par leur union. Lorsque ces deux fragments de matière opposés se combinent, ils se détruisent (*annihilation mutuelle*), ne laissant subsister aucune matière ; il ne reste que de l'énergie, sous forme de rayons gamma. Ainsi était confirmée l'idée d'Albert Einstein, selon laquelle la matière pouvait être convertie en énergie et vice-versa. Effectivement, Anderson parvint rapidement à détecter le processus inverse : des rayons gamma disparaissent soudainement en donnant naissance à des paires électron-positron. C'est ce que l'on appelle la *production de paire.* (Le prix Nobel de physique 1936 a été décerné à Anderson et Hess.)

Peu après, les Joliot-Curie tombèrent sur le positron dans une autre circonstance et firent ainsi une découverte fondamentale. En bombardant des atomes d'aluminium avec des particules alpha, ils trouvèrent que ce procédé produisait non seulement des protons, mais aussi des positrons. Le fait en lui-même était intéressant mais n'avait rien de renversant. Cependant, après arrêt du bombardement, l'aluminium continuait d'émettre des positrons. L'émission s'atténuait avec le temps. Apparemment ils avaient réussi à créer un nouveau corps radioactif dans la cible.

Les Joliot-Curie interprétèrent ce qui s'était passé de la façon suivante : lorsqu'un noyau d'aluminium absorbait une particule alpha, l'addition de deux protons changeait l'aluminium (numéro atomique 13) en phosphore (numéro 15). Comme la particule alpha contenait au total quatre nucléons, le nombre de masse devait s'élever de quatre unités, de l'aluminium 27

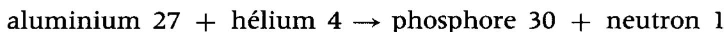
au phosphore 31. Alors, si la réaction expulsait un proton de ce noyau, la diminution de son numéro atomique et de son nombre de masse d'une unité devait le changer en un autre élément, à savoir le silicium 30.

Comme la particule alpha est un noyau d'hélium, et le proton un noyau d'hydrogène, on peut écrire l'équation suivante pour représenter cette *réaction nucléaire* :


aluminium 27 + hélium 4 → silicium 30 + hydrogène 1

Remarquez l'équilibre des nombres de masse : 27 + 4 = 30 + 1. De même pour les numéros atomiques, le 13 de l'aluminium et le 2 de l'hélium font 15, tout comme le 14 du silicium et le 1 de l'hydrogène. Cet équilibre des nombres de masse et des numéros atomiques est une règle générale des réactions nucléaires.

Les Joliot-Curie supposèrent que des neutrons pouvaient se former aussi bien que des protons au cours de la réaction. Si le phosphore 31 émettait un neutron au lieu d'un proton, le numéro atomique ne changeait pas, même si le nombre de masse diminuait d'une unité. Dans ce cas, l'élément était toujours du phosphore, mais du phosphore 30. Cette réaction avait pour équation :


aluminium 27 + hélium 4 → phosphore 30 + neutron 1

Comme le numéro atomique du phosphore est 15 et celui du neutron 0, les numéros atomiques sont encore bien équilibrés dans cette équation.

Les deux processus – l'absorption d'un alpha suivie de l'émission d'un neutron – se produisent lorsque l'aluminium est bombardé par des particules alpha. Mais il y a une distinction importante entre les deux résultats. Le silicium 30 est un isotope du silicium bien connu, qui représente un peu plus de 3 % du silicium naturel. Par contre le phosphore 30 n'existe pas dans la nature. La seule forme naturelle de phosphore est le phosphore 31. En bref, le phosphore 30 est un isotope radioactif de courte durée de vie, qui n'existe aujourd'hui que dans la mesure où il est produit artificiellement ; ce fut en fait le premier isotope de ce type fabriqué en laboratoire. Les Joliot-Curie ont reçu le prix Nobel de physique en 1935 pour leur découverte de la radioactivité artificielle.

Le phosphore 30 instable que les Joliot-Curie avaient produit par bombardement de l'aluminium se désintégrait rapidement en émettant des positrons. Puisque le positron, comme l'électron, est pratiquement sans masse, cette émission ne changeait pas le nombre de masse du noyau. Toutefois, la perte d'une charge positive abaissait son numéro atomique d'une unité, le métamorphosant de phosphore en silicium.

D'où vient le positron ? Les positrons figurent-ils au nombre des constituants du noyau ? La réponse est non. Simplement, un proton dans le noyau se change en neutron en abandonnant sa charge positive, libérée sous forme d'un positron rapide.

On peut maintenant expliquer l'émission de particules bêta, l'énigme que nous avons rencontrée plus tôt dans ce chapitre. Cette émission résulte du processus exactement inverse de la désintégration d'un proton en un neutron : à savoir, un neutron se change en proton. La métamorphose du proton en neutron libère un positron et, par symétrie, la conversion du neutron en proton produit un électron (la particule bêta). La perte d'une charge négative est équivalente au gain d'une charge positive et rend compte de la formation d'un proton chargé positivement à partir d'un

neutron non chargé. Mais comment le neutron sans charge parvient-il à s'extraire une charge négative et à l'éjecter à grande vitesse ?

Cela serait en fait impossible au neutron s'il se contentait de ne produire qu'une charge négative. Deux siècles d'expériences ont enseigné aux physiciens qu'une charge négative, pas plus qu'une charge positive, ne peut être créée à partir de rien. Pas plus qu'une de ces charges, quel que soit son signe, ne peut être détruite. C'est la loi de *conservation de la charge électrique.*

Mais en produisant une charge bêta un neutron ne crée pas seulement un électron ; il crée aussi un proton. Le neutron neutre disparaît pour laisser la place à un proton chargé positivement et un électron chargé négativement. *Prises ensemble,* les deux particules créées ont une charge électrique globale nulle. On n'a créé aucune charge *nette.* De même, lorsqu'un positron et un électron se rencontrent et se condamnent à l'annihilation mutuelle, d'entrée la charge du positron et de l'électron *pris ensemble* est nulle.

Lorsqu'un proton émet un positron et se change en neutron, la particule initiale (le proton) a une charge positive et les particules finales (le neutron et le positron), prises ensemble, ont une charge positive.

Un noyau a aussi la faculté d'absorber un électron. Lorsque cela se produit, un proton dans le noyau se change en neutron. Un électron plus un proton (qui pris ensemble ont une charge nulle) forment un neutron qui a une charge nulle. L'électron capturé est sur la couche électronique de l'atome la plus intérieure, car ce sont les électrons de cette couche qui sont les plus proches du noyau et donc les plus faciles à attirer. Ce processus est appelé *capture K* parce que la couche la plus intérieure est la couche K (voir le chapitre 6). Un électron de la couche L tombe alors dans la place vacante en émettant un rayonnement X. C'est par ces rayons X que l'on peut détecter la capture K. Cette détection a été réalisée pour la première fois par le physicien américain Luis Walter Alvarez en 1938. Les réactions nucléaires ne sont ordinairement pas affectées par les modifications chimiques qui ne touchent qu'aux électrons. Comme la capture K met en jeu tout autant les électrons que les noyaux, les chances pour qu'elle se produise peuvent être quelque peu modifiées à la suite de transformations chimiques.

Toutes ces interactions de particules satisfont la loi de conservation de la charge électrique et sont en outre soumises à d'autres lois de conservation.

Les physiciens croient que toute interaction de particules qui ne viole aucune loi de conservation peut finir par se produire, et qu'un observateur muni de la patience et des instruments adéquats la détectera. Les événements qui violent une loi de conservation sont « interdits » et ne peuvent se produire. Néanmoins, les physiciens ont parfois la surprise de trouver que ce qu'ils prenaient pour une loi de conservation n'est pas aussi rigoureux ou fondamental qu'ils le croyaient. C'est ce que nous verrons plus tard.

LES ÉLÉMENTS RADIOACTIFS

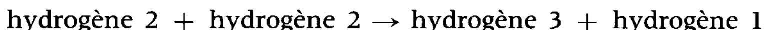
Dès que les Joliot-Curie eurent créé le premier isotope radioactif artificiel, les physiciens se mirent joyeusement à en produire des quantités. En fait, on a maintenant formé en laboratoire des variétés radioactives de tous

les éléments du tableau périodique. Dans le tableau périodique actuel, chaque élément est en réalité une famille, avec des membres stables et des membres instables, dont certains existent dans la nature et les autres seulement en laboratoire.

L'hydrogène, par exemple, existe sous trois formes. Il y a d'abord l'hydrogène ordinaire qui ne contient qu'un proton solitaire. Le chimiste Harold Urey est parvenu en 1932 à isoler une deuxième forme en évaporant une grande quantité d'eau, dans l'espoir qu'il ne resterait à la fin qu'un concentré de la forme la plus lourde de l'hydrogène, dont il soupçonnait l'existence. Comme prévu, lorsqu'il étudia la spectroscopie des dernières gouttes d'eau non évaporée, il observa une faible raie dans le spectre, exactement à l'endroit prévu pour *l'hydrogène lourd.*

Le noyau d'hydrogène lourd est constitué d'un proton et d'un neutron. Son nombre de masse valant deux, cet isotope est l'hydrogène 2. Urey a nommé cet atome *deutérium,* d'après un mot grec signifiant « deuxième », et son noyau le *deutéron,* devenu *deuton* en français. On appelle *eau lourde* une molécule d'eau à base de deutérium. Comme la masse du deutérium est le double de celle de l'hydrogène ordinaire, l'eau lourde a des points d'ébullition et de fusion plus élevés que ceux de l'eau ordinaire. Alors que l'eau ordinaire bout à 100 °C et gèle à 0 °C, l'eau lourde bout à 101,42 °C et gèle à 3,79 °C. Le deutérium lui-même bout à 23,7 °K, comparé à 20,4 °K pour l'hydrogène ordinaire. La découverte du deutérium a valu à Urey le prix Nobel de chimie en 1934.

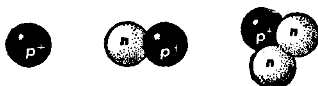
Le deuton s'est avéré d'un intérêt inestimable pour bombarder les noyaux. En 1934, le physicien australien Marcus Lawrence Elwin Oliphant et le chimiste autrichien Paul Harteck ont attaqué le deutérium lui-même avec des deutons. Il s'est produit la réaction suivante :


hydrogène 2 + hydrogène 2 → hydrogène 3 + hydrogène 1

Le nouvel hydrogène « superlourd » a été baptisé *tritium,* d'après le mot grec pour « troisième », et son noyau le *triton.* Il bout à 25,0 °K et fond à 20,5 °K. On a pu préparer de l'oxyde de tritium pur (*l'eau superlourde*), dont le point de fusion est à 4,5 °C. Le tritium est radioactif et se désintègre relativement vite. On le trouve dans la nature, formé dans les produits du bombardement de l'atmosphère par le rayonnement cosmique. En se désintégrant, il émet un électron et se change en hélium 3, un isotope stable, quoique rare, de l'hélium, mentionné dans le chapitre précédent (figure 7.3).

Dans l'hélium atmosphérique, il n'y a environ qu'un atome d'hélium 3 sur 800 000, provenant sans nul doute de la désintégration de l'hydrogène 3 (le tritium), lui-même formé dans les réactions nucléaires qui se produisent lorsque les particules des rayons cosmiques frappent les atomes de l'atmosphère. Le tritium présent à un instant donné est encore plus rare.

Figure 7.3. Les noyaux d'hydrogène ordinaire, de deutérium et de tritium.



On estime qu'il n'y en a au total qu'un kilo et demi dans l'atmosphère et l'ensemble des océans. Le pourcentage d'hélium 3 dans l'hélium des gisements de gaz naturel, où les rayons cosmiques ont moins d'occasions de former du tritium, est encore plus faible.

L'hélium 3 et l'hélium 4 ne sont pas les deux seuls isotopes de l'hélium. Les physiciens ont créé deux formes radioactives : l'hélium 5, l'un des noyaux les plus instables connus, et l'hélium 6, lui aussi très instable. Et cela continue. Il y a maintenant environ 1 400 isotopes connus. Plus de 1 100 d'entre eux sont radioactifs, parmi lesquels un grand nombre ont été créés par de nouveaux procédés d'artillerie atomique beaucoup plus puissants que les particules alpha des sources radioactives, les seuls projectiles dont disposaient Rutherford et les Joliot-Curie.

Le genre d'expériences réalisées par les Joliot-Curie au début des années trente qui semblait, à l'époque, réservé à la tour d'ivoire des scientifiques, a finalement une application éminemment pratique. Bombardons avec des neutrons un ensemble d'atomes identiques ou de différentes sortes. Un certain pourcentage d'atomes de chaque sorte vont absorber un neutron, dont il résultera généralement des atomes radioactifs. Ces éléments radioactifs vont se désintégrer, en émettant des rayonnements nucléaires sous forme de particules ou de rayons gamma.

Chaque type différent d'atome absorbera un neutron pour former un type d'atome radioactif qui émettra un rayonnement caractéristique. On peut détecter ce rayonnement avec une grande précision. A partir de son type et de la décroissance de son taux de production, on peut identifier l'atome radioactif qui l'a émis, et donc l'atome initial avant absorption d'un neutron. On peut analyser les corps de cette façon (*analyse par activation*) avec une précision sans précédent : on peut détecter des quantités d'un nuclide particulier qui n'excèdent pas le picogramme.

On peut utiliser l'analyse par activation pour mettre en évidence des différences imperceptibles dans les impuretés contenues dans des échantillons de pigments en usage à différentes époques, et parvenir de cette façon à déterminer l'authenticité d'un tableau apparemment ancien en ne prélevant qu'un minuscule fragment de peinture. Le procédé permet d'autres décisions délicates du même genre ; on a même étudié un cheveu prélevé sur le corps de Napoléon, vieux d'un siècle et demi, et trouvé qu'il contenait de l'arsenic, mais il est difficile de dire si cet arsenic provient d'un empoisonnement criminel ou de médicaments, ou si sa présence n'est que fortuite.

LES ACCÉLÉRATEURS DE PARTICULES

Dirac avait annoncé l'existence non seulement d'un antiélectron (le positron) mais aussi d'un antiproton. Mais, pour produire un antiproton il fallait beaucoup plus d'énergie que pour un positron. L'énergie nécessaire est proportionnelle à la masse de la particule. Comme la masse du proton est 1 836 fois plus élevée que celle de l'électron, la formation d'un antiproton exige, au minimum, 1 836 fois plus d'énergie que celle d'un positron. Cette réalisation devait attendre la mise au point d'un appareil capable d'accélérer les particules subatomiques à des énergies suffisamment élevées.

A l'époque, on venait juste de commencer des recherches dans cette direction. En 1928, les physiciens anglais John Douglas Cockcroft et Ernest

Thomas Sinton Walton, qui travaillaient dans le laboratoire de Rutherford, réalisèrent un *multiplicateur de tension,* un système pour créer un potentiel élevé qui pouvait propulser un proton jusqu'à une énergie proche de 400 000 électron-volts. (Un *électron-volt* représente l'énergie d'un électron qui a été accéléré par le champ électrique créé par une différence de potentiel de 1 volt.) Avec les protons accélérés dans cette machine, Cockcroft et Walton pouvaient briser le noyau de lithium, succès qui leur valut le prix Nobel de physique en 1951.

Dans le même temps, le physicien américain Robert Jemison Van de Graaff créait un autre type d'accélérateur. Pour l'essentiel, cette machine séparait les électrons des protons et, au moyen d'une courroie de transport, les accumulait en ses deux extrémités opposées. Le *générateur électrostatique de Van de Graaff* réalisait de cette façon une différence de potentiel électrique très élevée entre les deux extrémités ; Van de Graaff parvint jusqu'à 8 millions de volts. Les générateurs électrostatiques peuvent maintenant accélérer facilement les protons jusqu'à des vitesses correspondant à 24 millions d'électron-volts (les physiciens emploient l'abréviation *MeV* pour million d'électron-volts).

Les photographies spectaculaires du générateur de Van de Graaff en train de produire d'énormes étincelles captivèrent l'imagination populaire et initièrent le public à ce « briseur d'atome ». On le considérait généralement comme un système pour produire des éclairs « de la main de l'homme », bien qu'il représentât en fait beaucoup plus. (Un générateur conçu pour la production d'éclairs artificiels, et rien de plus, avait déjà été réalisé en 1922 par l'ingénieur électricien américain d'origine allemande, Charles Proteus Steinmetz.)

Du point de vue pratique, l'énergie atteinte dans ce genre de machine est bornée par les limites imposées au potentiel réalisable. Mais un autre principe pour accélérer les particules vit bientôt le jour. Imaginons qu'au lieu de projeter les particules d'un seul coup violent, on les accélère par une série de petites poussées. Si chacune des poussées successives est exercée juste au bon moment, elle permet d'augmenter à chaque fois la vitesse, comme les poussées sur une balançoire propulsent celle-ci de plus en plus haut, si elles sont appliquées « en phase » avec les oscillations de la balançoire.

Cette idée a donné naissance en 1931 à l'*accélérateur linéaire* (figure 7.4). Les particules sont expédiées dans un tube divisé en sections. La force propulsive est assurée par un champ électrique alternatif conçu pour que les particules reçoivent une poussée supplémentaire à chaque entrée dans une nouvelle section où elles vont dériver. Comme les particules vont de plus en plus vite, chaque tube de dérive doit être un peu plus long que le précédent, de sorte que les particules y passent toujours le même temps et restent en phase avec les poussées.

La synchronisation n'est pas facile à maintenir, et il y a de toute façon une limite à la longueur du tube pratiquement réalisable, ce qui explique l'oubli relatif dont fut victime l'accélérateur linéaire durant les années trente. Une des raisons qui le reléguèrent à l'arrière-plan tenait au fait qu'Ernest Orlando Lawrence, de l'université de Californie, avait eu une meilleure idée.

Au lieu de propulser les particules dans un tube en ligne droite, pourquoi ne pas les faire tourner suivant une trajectoire circulaire ? C'était réalisable à l'aide d'un aimant. Après chaque demi-tour on donnait une chiquenaude

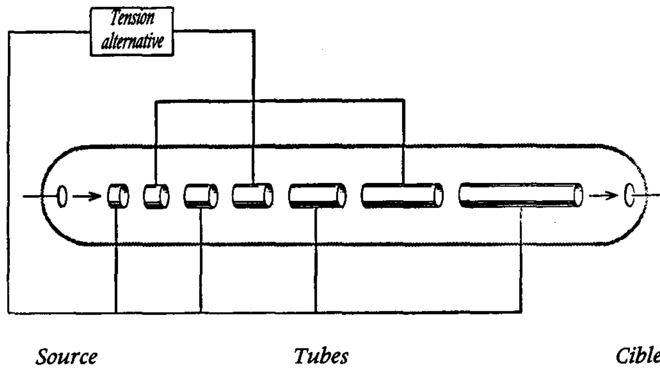


Figure 7.4. Principe de l'accélérateur linéaire. Une alimentation alternative à haute fréquence pousse et tire alternativement les particules chargées entre les sections de tube successives, ce qui les accélère, ici vers la droite.

aux particules au moyen du champ alternatif et, avec cette disposition, il n'était plus aussi difficile de contrôler le synchronisme. A mesure que les particules allaient de plus en plus vite, leur trajectoire était de moins en moins courbée par le champ magnétique, de sorte qu'elles se déplaçaient sur des arcs de plus en plus grands et mettaient le même temps pour chaque tour. A la fin de leur course en spirale, les particules émergeaient de la chambre circulaire (en fait divisée en demi-cercles, appelés *dés* à cause de leur forme qui, dans son principe tout au moins, rappelle celle de la lettre D ; voir la figure 7.5) pour aller frapper leur cible.

Le nouveau système compact de Lawrence a été appelé *cyclotron* (figure 7.5). Son premier modèle, qui avait moins de trente centimètres de diamètre, pouvait accélérer des protons jusqu'à des énergies de près de 1,25 MeV. Dès 1939, l'université de Californie avait un cyclotron dont les aimants, larges de 1,50 mètre, permettaient d'accélérer les particules jusqu'à 20 MeV, soit à une vitesse double de celle des particules alpha les plus énergétiques émises par les sources radioactives. Cette année-là, Lawrence reçut le prix Nobel de physique pour son invention.

Le cyclotron lui-même devait se limiter à une vingtaine de MeV, car au-delà de cette énergie, les particules vont si vite que l'augmentation de la masse avec la vitesse (un effet prédit par la théorie de la relativité d'Einstein) devient appréciable. Du fait de cette augmentation de leur inertie, les particules commencent à traîner et à prendre du retard sur les poussées du champ électrique. Mais il y avait un remède à cela, qui fut mis au point en 1945 par le physicien soviétique Vladimir Iosifovitch Veksler et, indépendamment, par le physicien californien Edwin Mattison McMillan. La solution consistait simplement à adapter la fréquence des oscillations du champ électrique à l'augmentation de la masse des particules. Ce cyclotron modifié a été appelé *synchrocyclotron*. En 1946, l'université de Californie en avait construit un qui pouvait accélérer les particules à des énergies de 200 à 400 MeV. Des synchrocyclotrons plus grands, construits plus tard aux États-Unis et en Union Soviétique, atteignent des énergies de 700 à 800 MeV.

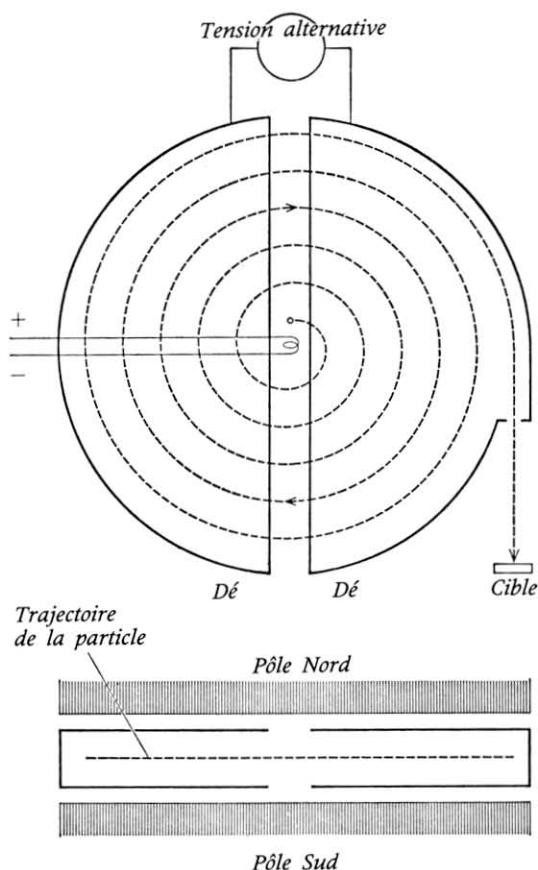


Figure 7.5. Schéma de principe du cyclotron, vu de dessus *(en haut)* et vu de côté *(en bas)*. Les particules, injectées par la source, reçoivent une poussée à chaque passage entre les dés grâce à l'alimentation alternative, et sont défléchies sur une trajectoire spirale par un aimant.

Entre-temps, l'accélération des électrons avait été de son côté l'objet de soins attentifs. Pour qu'ils puissent briser les atomes, il faut donner aux électrons, qui sont légers, des vitesses beaucoup plus élevées qu'aux protons (comme une balle de ping-pong doit aller beaucoup plus vite qu'une balle de golf pour causer les mêmes dégâts). Le cyclotron ne pouvait convenir pour des électrons, car leur augmentation de masse devenait trop grande aux vitesses élevées nécessaires pour qu'ils soient efficaces. En 1940, le physicien américain Donald William Kerst conçut un accélérateur d'électrons qui compensait l'augmentation de masse par un champ électrique d'intensité croissante. Les électrons restaient sur la même orbite circulaire au lieu de décrire une spirale vers l'extérieur. Cet instrument fut baptisé le *bêtatron* d'après les particules bêta. Les bêtatrons permettent maintenant d'accélérer des électrons jusqu'à 340 MeV.

Ils ont été rejoints par un autre instrument, de conception légèrement différente, appelé *synchrotron à électrons*. Le premier de ces synchrotrons a été construit en Angleterre, en 1946, par F. K. Goward et D. E. Barnes. Ces appareils peuvent élever l'énergie des électrons jusqu'à 1 000 MeV, valeur qu'ils ne peuvent dépasser, car des électrons sur une trajectoire circulaire rayonnent de l'énergie à un taux croissant lorsque la vitesse augmente. Ce rayonnement produit par une particule chargée lorsqu'elle est déviée est appelé *Bremsstrahlung* (ou rayonnement de freinage), un mot allemand qui signifie « rayonnement de freinage ».

Attaquant un nouveau chapitre après le bêtatron et le synchrotron à électrons, les physiciens qui utilisaient des protrons ont commencé vers 1947 à construire des *synchrotrons à protons* qui, de la même façon, gardent leurs particules sur une orbite circulaire unique. Ceci permet une économie de poids. Quand les particules se déplacent vers l'extérieur, sur une spirale, l'aimant doit couvrir toute la largeur de cette spirale pour que s'exerce partout sa force magnétique. Avec une seule trajectoire circulaire, il suffit d'un aimant qui couvre une zone étroite.

Comme le protron, plus massif, ne perd pas son énergie aussi rapidement que l'électron lorsqu'il se meut sur une trajectoire circulaire, les physiciens décidèrent de franchir la limite des 1 000 MeV avec un synchrotron à protons. Cette valeur de 1 000 MeV correspond à un milliard d'électron-volts, ou un giga-électron-volt, abrégé en *GeV.*

En 1952 fut terminé, au laboratoire national de Brookhaven, à Long Island, un synchrotron à protons qui atteignait 2 à 3 GeV. On l'a appelé le *cosmotron,* car il se trouve dans le domaine d'énergie des particules des rayons cosmiques. Deux ans plus tard, l'université de Californie terminait son *bévatron,* capable de produire des particules entre 5 et 6 GeV. Puis en 1957, l'Union Soviétique annonçait que son *phasotron* était parvenu à 10 GeV.

Mais ces machines sont maintenant dépassées par les accélérateurs d'un nouveau type appelé *synchrotron à focalisation forte.* La limitation du type bévatron vient de ce que les particules du faisceau s'écartent et vont se perdre dans les parois du tube qui fait office de chambre à vide dans laquelle se déplace le faisceau. Le nouveau système s'oppose à cette tendance au moyen de champs magnétiques alternés de formes différentes qui maintiennent les particules focalisées en un faisceau fin. Cette idée a été suggérée par Christofilos, dont les talents « d'amateur » ont ici surclassé les professionnels, comme dans le cas de l'effet Christofilos. Par ailleurs, ce système a encore diminué le volume de l'aimant nécessaire aux énergies atteintes. En augmentant l'énergie d'un facteur cinquante, l'aimant n'a même pas doublé.

En novembre 1959, le Comité Européen pour la Recherche Nucléaire (le CERN), une organisation à laquelle coopèrent douze nations, a terminé à Genève un synchrotron à focalisation forte qui atteint 24 GeV et produit de grosses bouffées de particules (contenant 10 milliards de protons) toutes les trois secondes. Ce synchrotron a un diamètre de près de 200 mètres et une circonférence de 600 mètres. Durant les trois secondes de formation et d'accélération de la bouffée, les protons décrivent un demi-million de tours de piste. L'instrument a coûté 30 millions de dollars et son aimant pèse 3 500 tonnes.

On n'arrête pas le progrès. Il fallait des énergies de plus en plus élevées pour produire des réactions de plus en plus inhabituelles entre les

particules, pour former des particules de plus en plus massives, et pour en apprendre de plus en plus sur la structure ultime de la matière. Par exemple, au lieu d'accélérer un faisceau de particules que l'on envoie taper sur une cible fixe, pourquoi ne pas préparer deux faisceaux de particules, orbitant en sens opposé dans des *anneaux de stockage* où l'on se contente d'entretenir leur vitesse pendant un certain temps ? A des moments choisis, les deux faisceaux sont dirigés de façon à entrer en collision face à face. L'énergie effective dans ce type de collision est quatre fois plus grande que si un des faisceaux frappait une cible de même nature mais fixe. Au Fermilab (Fermi National Accelerator Laboratory), près de Chicago, un accélérateur fonctionnant sur ce principe est entré en service en 1982 et devrait atteindre 1 000 GeV. On l'appelle le *tévatron* (*t* pour « trillion », qui signifie en anglais mille milliards). On prévoit d'autres accélérateurs qui pourraient atteindre 20 000 GeV.

Parallèlement on a assisté à une renaissance de l'accélérateur linéaire, ou *linac,* grâce à des améliorations techniques permettant d'éliminer les inconvénients des premiers modèles. Aux énergies très élevées, un accélérateur linéaire présente certains avantages sur le type circulaire. Comme les électrons perdent moins d'énergie par rayonnement lorsqu'ils sont accélérés en ligne droite, un linac permet d'accélérer les électrons à des vitesses plus élevées, et de mieux focaliser les faisceaux sur les cibles. L'université de Stanford a construit un accélérateur linéaire long de trois kilomètres qui peut atteindre des énergies de 45 GeV.

Rien qu'avec le bévatron, l'humanité a accédé enfin au pouvoir de créer l'antiproton. Les physiciens californiens s'étaient fixé pour objectif de le produire et de le détecter. En 1955, après avoir bombardé du cuivre avec des protons de 6,2 GeV heure après heure, Owen Chamberlain et Emilio G. Segrè ont capturé sans conteste l'antiproton (en fait soixante antiprotons). Leur identification était loin d'être facile. Pour chaque antiproton produit il y avait 40 000 particules d'autres types créées. Mais grâce à un système de détecteurs perfectionné, conçu pour que seul un antiproton satisfasse tous les critères de déclenchement, Chamberlain et Segrè ont pu identifier cette particule sans aucune ambiguïté. Ce succès leur a valu le prix Nobel de physique en 1959.

L'antiproton est aussi éphémère que le positron, tout au moins dans notre univers. Une fraction de seconde après sa création, cette particule est happée par un noyau normal chargé positivement. L'antiproton et l'un des protons du noyau s'annihilent mutuellement, pour donner de l'énergie et des particules plus légères. En 1965, on est parvenu à concentrer suffisamment d'énergie au départ pour inverser le processus et produire une paire proton-antiproton.

De temps en temps, la collision du proton et de l'antiproton n'est que tangentielle plutôt que directe. Lorsque ceci se produit, leurs charges respectives se neutralisent. Le proton est changé en neutron, ce qui est son droit le plus strict. Mais l'antiproton devient un *antineutron !* Comment peut-on être antineutron ? Le positron est l'opposé de l'électron en vertu de sa charge électrique opposée, et de même, l'antiproton est « anti » par sa charge. Mais qu'est-ce qui rend « anti » cet antineutron qui n'est pas chargé ?

LE SPIN DES PARTICULES

Arrivés là, il nous faut à nouveau reparler du spin d'une particule, une propriété d'abord suggérée en 1925 par les physiciens hollandais George Eugene Uhlenbeck et Samuel Abraham Goudsmit. Comme si elle tournait sur elle-même, la particule génère un minuscule champ magnétique ; de tels champs magnétiques ont été mesurés et soigneusement étudiés, en particulier par le physicien allemand Otto Stern et le physicien américain Isidor Rabi, qui furent respectivement lauréats des prix Nobel de physique 1943 et 1944 pour leurs contributions dans ce domaine.

Les particules – comme le proton, le neutron et l'électron –, dont les spins ont des valeurs demi-entières, doivent être traitées selon un système de règles établies indépendamment en 1926 par Fermi et Dirac. C'est la raison pour laquelle on appelle *statistique de Fermi-Dirac* cet ensemble de règles. Les particules qui y obéissent sont des *fermions ;* le proton, l'électron et le neutron sont donc tous des fermions.

Il existe aussi des particules dont le spin se mesure par des nombres entiers. Il faut les traiter selon un autre ensemble de règles, édictées par Einstein et le physicien indien Satyendranath Bose. Les particules qui se comportent selon la *statistique de Bose-Einstein* sont des *bosons.* La particule alpha, par exemple, est un boson.

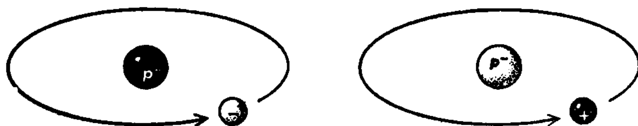
Ces deux classes de particules ont des propriétés différentes. Le principe d'exclusion de Pauli (voir le chapitre 6) s'applique non seulement aux électrons, mais aussi à toutes les espèces de fermions. En revanche, il ne concerne pas les bosons.

On comprend facilement qu'une particule chargée crée un champ magnétique, mais il est plus difficile d'imaginer qu'un neutron non chargé puisse en faire autant. Et pourtant il le fait. La preuve la plus directe en est qu'un faisceau de neutrons frappant du fer a un comportement différent selon que le fer est magnétisé ou non. Le magnétisme du neutron vient certainement du fait que (comme nous le verrons) cette particule est constituée d'autres particules qui, elles, sont porteuses de charge électrique. Regroupées dans le neutron, ces charges se compensent, mais elles se débrouillent quand même pour établir un champ magnétique.

De toute façon, le spin du neutron nous apporte la réponse à la question de la nature de l'antineutron. Ce n'est qu'un neutron dont le sens du spin est inversé ; son pôle sud magnétique est, disons, en haut plutôt qu'en bas. En fait, le proton et l'antiproton présentent exactement le même phénomène d'inversion des pôles que l'électron et le positron.

Les antiparticules peuvent, sans aucun doute, se combiner pour former de l'*antimatière,* comme les particules ordinaires forment la matière (figure 7.6). Le premier échantillon d'antimatière a été fabriqué à Brookhaven

Figure 7.6. Un atome d'hydrogène et l'atome d'antimatière correspondant, formé d'un antiproton et d'un positron.



en 1965. Le bombardement d'une cible de béryllium avec des protons de 7 GeV y a produit des associations d'antiprotons et d'antineutrons, quelque chose comme des *antideutons*. On a depuis produit de l'*antihélium 3* et on peut certainement, en s'en donnant la peine, former des antinoyaux encore plus complexes. Au moins le principe est clair, et nul physicien n'en doute. L'antimatière peut exister.

Mais existe-t-elle en réalité ? Y a-t-il des masses d'antimatière dans l'Univers ? S'il y en a, elles ne se trahissent pas à distance. Leurs effets gravitationnels et la lumière qu'elles émettent sont absolument identiques à ceux de la matière ordinaire. Mais si elles rencontrent de la matière ordinaire, les formidables réactions d'annihilation qui en résultent doivent se remarquer. Elles le doivent peut-être, mais on ne les a pas remarquées. Se peut-il donc que l'Univers soit principalement constitué de matière, avec peu ou pas du tout d'antimatière ? Si tel est le cas, pourquoi ? Puisque matière et antimatière sont équivalentes en tout point, excepté leur symétrie électromagnétique, toute force qui crée l'une devrait créer l'autre, et l'Univers devrait être constitué de quantités égales de chacune.

Voilà un dilemme. La théorie nous dit qu'il devrait y avoir de l'antimatière là dehors, dans l'espace, ce que l'observation refuse de confirmer. Mais pouvons-nous être certains de ne rien observer ? Qu'en est-il du cœur des galaxies actives et, a fortiori, des quasars ? Ces phénomènes très énergétiques pourraient-ils résulter de l'annihilation matière-antimatière ? Probablement non ! Même cette annihilation semble insuffisante, et les astronomes envisagent plutôt, comme seul mécanisme qui produirait l'énergie nécessaire, l'effondrement gravitationnel et le phénomène de trou noir.

LES RAYONS COSMIQUES

Qu'en est-il, dans ces conditions, des rayons cosmiques ? La plupart des particules des rayons cosmiques ont une énergie comprise entre 1 et 10 GeV. On peut en rendre compte par l'interaction matière-antimatière ; mais il y a des particules cosmiques qui vont beaucoup plus vite : 20 GeV, 30 GeV, 40 GeV (voir la figure 7.7). Les physiciens de l'Institut de technologie du Massachusetts ont même détecté quelques particules qui avaient une énergie colossale, de 20 milliards de GeV. Des nombres comme celui-là dépassent ce que peut appréhender l'esprit, mais on peut avoir une idée de ce que signifie cette énergie lorsqu'on réalise qu'avec 20 milliards de GeV, une particule submicroscopique pourrait soulever de 15 centimètres un poids de 2 kilos.

Depuis la découverte des rayons cosmiques, les gens n'ont cessé de se demander d'où ils venaient et comment ils avaient été créés. L'hypothèse la plus simple est que quelque part dans la Galaxie, peut-être dans notre Soleil, peut-être plus loin, se produisent des réactions nucléaires qui expulsent des particules avec les énergies énormes que nous leur trouvons. Effectivement, des bouffées de rayons cosmiques modérés nous parviennent environ tous les deux ans, en relation avec les éruptions solaires, comme on l'a observé pour la première fois en 1942. Que devrait-il en être alors de sources comme les supernovae, les pulsars et les quasars ? Mais aucune réaction nucléaire connue ne peut produire quoi que ce soit avec une énergie de 20 milliards de GeV. L'annihilation mutuelle des

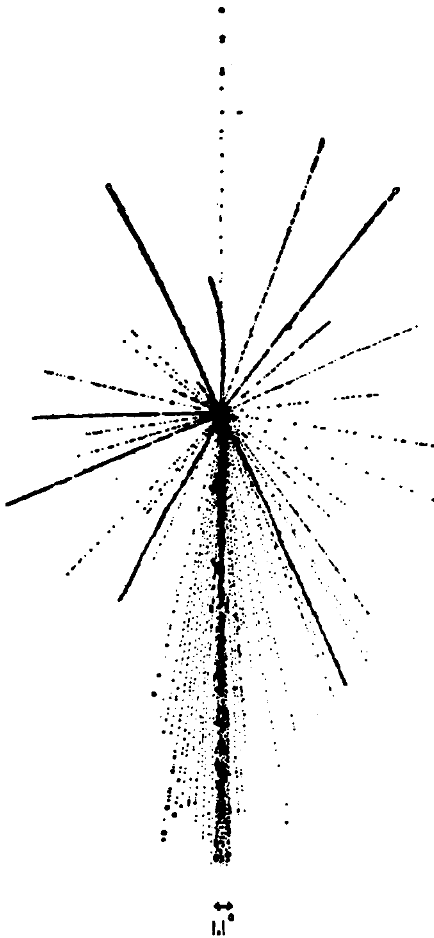


Figure 7.7. Un rayon cosmique de 30 000 Gev fracasse un atome d'argent. La collision de la particule cosmique avec le noyau d'argent a produit 95 fragments nucléaires, dont les traces forment une étoile.

noyaux de matière et d'antimatière les plus lourds ne produirait que des particules avec une énergie maximale de 250 GeV.

L'alternative consiste à supposer, comme l'a fait Fermi, qu'il y a dans l'espace une force qui accélère les particules cosmiques. Les particules peuvent naître avec une énergie modérée à la faveur d'explosions comme celles des supernovae, et se trouver progressivement accélérées dans leur voyage à travers l'espace.

Selon la théorie la plus en vogue actuellement, les particules seraient accélérées par des champs magnétiques cosmiques agissant comme de gigantesques synchrotrons. Il y a effectivement des champs magnétiques dans l'espace, et on estime que notre galaxie doit posséder un tel champ, bien que son intensité n'excède pas 1/20 000 du champ magnétique terrestre.

En passant à travers ce champ, les particules cosmiques seraient doucement accélérées sur une trajectoire incurvée. En gagnant de l'énergie, leurs trajectoires seraient de plus en plus larges, jusqu'à ce que les plus énergétiques s'échappent de la Galaxie. Bien que la plupart des particules n'atteignent jamais cette trajectoire de libération, car elles perdent de l'énergie par collision avec d'autres particules ou avec des corps plus importants, c'est possible pour quelques-unes. Effectivement, les particules cosmiques les plus énergétiques qui nous parviennent pourraient bien avoir traversé notre galaxie après avoir été expulsées d'une autre galaxie par ce procédé.

LA STRUCTURE DU NOYAU

Après en avoir tant appris sur la nature et l'allure générale de la construction du noyau, on en vient maintenant à s'intéresser à sa structure, en particulier sa structure interne détaillée. Tout d'abord, quelle est sa forme ? Comme il est si petit, et rempli de neutrons et de protons de façon si serrée, les physiciens le supposent naturellement sphérique. Les détails fins des spectres des atomes suggèrent que la distribution de charge de nombreux noyaux est sphérique. Ce n'est pas le cas pour certains, qui se comportent comme s'ils avaient deux paires de pôles magnétiques, et on dit de ces noyaux qu'ils ont un *moment quadrupolaire.* Mais leur déviation par rapport à la forme sphérique est faible. Le cas le plus extrême est celui des noyaux des lanthanides, dont la distribution de charge semble être en sphéroïde allongé (en d'autres termes, comme un ballon de rugby). Même dans ce cas, le grand axe n'excède pas le petit axe de plus de 20 %.

Pour ce qui est de la structure interne du noyau, le modèle le plus simple représente le noyau comme un ensemble bien serré de particules, analogue à une goutte de liquide où les particules (les molécules) sont serrées, laissant peu d'espace interstitiel, dont la densité est pratiquement la même partout, et qui est limitée par une surface frontière bien définie.

Ce *modèle de la goutte liquide* a été élaboré pour la première fois en détail par Niels Bohr en 1936. Il propose une explication possible pour l'absorption et l'émission de particules par certains noyaux. Lorsqu'une particule pénètre dans le noyau, on peut penser que l'énergie de son mouvement se répartit sur chacune des particules étroitement serrées, de sorte qu'aucune des particules ne reçoit, de prime abord, suffisamment d'énergie pour s'échapper. Après peut-être un milliard de milliardième de seconde, durant lequel ont pu se produire des milliards de collisions aléatoires, certaines particules acquièrent assez d'énergie pour s'échapper du noyau.

Ce modèle peut également rendre compte de l'émission de particules alpha par les noyaux lourds. Ces gros noyaux peuvent vibrer comme le font les gouttes liquides, lorsque les particules qui les constituent se déplacent et échangent de l'énergie. Tous les noyaux pourraient vibrer ainsi, mais les plus gros noyaux sont moins stables et ont plus de chance d'éclater. Pour cette raison, des fragments du noyau composés des deux protons et deux neutrons d'une particule alpha (combinaison particulièrement stable) peuvent se séparer spontanément de la surface du noyau. Il en résulte un noyau plus petit, moins exposé à la rupture par vibration et qui peut être finalement stable.

Cette vibration peut aussi entraîner une autre sorte d'instabilité. Lorsqu'une grosse goutte de liquide en suspension dans un autre liquide est mise en oscillation par des courants dans le liquide ambiant, elle tend à se briser en sphères plus petites, généralement en moitiés à peu près égales. On a fini par découvrir en 1939 (une découverte que je décrirai plus complètement dans le chapitre 10) que l'on pouvait effectivement faire éclater certains gros noyaux de cette façon en les bombardant avec des neutrons. C'est ce que l'on appelle la *fission nucléaire.*

En fait, cette fission nucléaire peut parfois se produire sans intervention d'une particule pertubatrice extérieure. Le frémissement interne cause, de temps en temps, la séparation du noyau en deux. Les physiciens soviétiques ont effectivement détecté, en 1940, cette *fission spontanée* dans des atomes d'uranium. L'instabilité de l'uranium se traduit principalement par l'émission de particules alpha, mais il y a quand même, dans un demi-kilo d'uranium, quatre fissions spontanées par seconde, pour huit millions de noyaux qui émettent une particule alpha.

La fission spontanée se produit aussi dans le protoactinium, dans le thorium et, plus fréquemment, dans les éléments transuraniens. Plus les noyaux sont gros, plus la probabilité de fission spontanée augmente. Dans les éléments les plus lourds, celle-ci devient le mécanisme de désintégration prépondérant, surclassant de loin l'émission de particules alpha.

Un autre modèle du noyau très en vogue apparente le noyau à l'ensemble d'un atome en représentant les nucléons dans le noyau comme des électrons en orbite autour du noyau, qui occupent des couches et sous-couches, chacun n'ayant qu'une faible influence sur les autres. C'est ce que l'on appelle le *modèle en couches.*

On peut, par analogie avec les couches électroniques de l'atome, supposer que les noyaux, dont les couches nucléoniques extérieures sont pleines, devraient être plus stables que ceux dont les couches externes sont incomplètes. La théorie la plus simple prédisait que les noyaux comportant 2, 8, 20, 40, 70 ou 112 protons ou neutrons devraient être particulièrement stables. Mais ce n'est pas ce que l'on observe. La physicienne américaine d'origine allemande Maria Goeppert Mayer a montré comment les spins des protons et des neutrons pouvaient affecter la situation. Dans ces conditions les noyaux contenant 2, 8, 20, 50, 82 ou 126 protons ou neutrons doivent être particulièrement stables, ce qui correspond aux observations. Les noyaux avec 28 ou 40 protons ou neutrons doivent être assez stables. Tous les autres sont moins stables, sinon instables. Ces valeurs des capacités des couches sont généralement appelées *nombres magiques.*

Parmi les noyaux à nombres magiques, il y a l'hélium 4 (2 protons et 2 neutrons), l'oxygène 16 (8 protons et 8 neutrons), et le calcium 40 (20 protons et 20 neutrons), tous particulièrement stables et plus abondants dans l'Univers que les autres noyaux de taille voisine.

En ce qui concerne les nombres magiques plus élevés, on trouve l'étain avec dix isotopes stables ayant chacun 50 protons, et quatre isotopes du plomb avec 82 protons. Il existe cinq isotopes stables (correspondant chacun à un élément différent) avec 50 neutrons et sept isotopes stables avec 82 neutrons. C'est généralement près des nombres magiques que les prédictions détaillées du modèle en couches sont les mieux vérifiées. A mi-chemin entre les nombres magiques (comme dans le cas des lanthanides et des actinides), l'accord avec l'expérience est médiocre. Mais c'est justement dans ces régions intermédiaires que les noyaux sont plus éloignés

justement dans ces régions intermédiaires que les noyaux sont plus éloignés de la forme sphérique (présupposée par le modèle en couches), et ont une forme ellipsoïdale plus marquée. Le prix Nobel de physique 1963 a été attribué à Goeppert Mayer et à deux autres chercheurs, Wigner et le physicien allemand Johannes Hans Daniel Jensen, pour leurs contributions à la théorie.

Généralement, à mesure que les noyaux sont plus complexes, ils se font plus rares dans l'Univers ou sont moins stables, ou les deux. Les isotopes stables les plus complexes sont le plomb 208 et le bismuth 209, chacun avec un nombre magique de 126 neutrons et, de surcroît, pour le plomb, le nombre magique de 82 protons. Au-delà, tous les nuclides sont instables et le sont en général de plus en plus lorsque la taille du noyau augmente. Mais les considérations de nombres magiques expliquent le fait que le thorium et l'uranium possèdent des isotopes beaucoup moins instables que les autres nuclides de taille similaire. Comme je l'ai déjà mentionné, la théorie prédit également que certains isotopes des éléments 110 et 114 pourraient être beaucoup moins instables que les autres nuclides de cette taille. Pour ceux-ci, il nous faut attendre de voir.

Les leptons

L'électron et le positron sont remarquables par la petitesse de leur masse, seulement 1/1836 fois celle du proton, du neutron, de l'antiproton ou de l'antineutron, et ils sont donc désignés globalement sous le nom de *leptons* (du grec signifiant « ténu »).

Bien qu'il se soit écoulé près d'un siècle depuis la découverte de l'électron, on n'a pas encore trouvé de particule qui soit à la fois plus légère que l'électon (ou le positron) et chargée électriquement. On ne s'attend d'ailleurs pas à ce genre de découverte. Il se peut que, à la charge électrique, quelle qu'en soit la nature (nous savons ce qu'elle fait et comment mesurer ses propriétés, mais nous ignorons ce qu'elle *est*), soit toujours asssociée une masse minimale, et que ce soit celle-ci qui se manifeste dans l'électron. Peut-être qu'en fait l'électron n'est rien d'autre que sa charge ; lorsque l'électron se comporte comme une particule, la charge électrique sur cette particule semble n'avoir aucune extension et n'occuper qu'un point.

Il existe bien des particules dont la masse est nulle (en fait, la *masse au repos* que j'expliquerai dans le prochain chapitre), mais elles n'ont pas de charge électrique. Par exemple, les ondes lumineuses et les autres formes d'ondes électromagnétiques peuvent se comporter comme des particules (voir le chapitre suivant). Cette manifestation « particulière » de ce que l'on se représente habituellement comme une onde est appelée *photon,* d'après le mot grec qui signifie « lumière ».

Le photon a une masse et une charge électrique nulles, mais son spin vaut 1 ; c'est donc un boson. Comment peut-on connaître la valeur de son spin ? Les photons participent aux réactions nucléaires, ils peuvent être absorbés dans certains cas, émis dans d'autres. Dans ces réactions nucléaires, le spin total des particules participantes, avant et après la

réaction, doit rester inchangé (c'est la *conservation du spin**). La seule façon d'assurer cette conservation dans les réactions nucléaires mettant en jeu des photons consiste à supposer que le photon a un spin égal à 1. Le photon n'est pas rangé dans la catégorie des leptons, car on réserve ce terme aux fermions.

On a des raisons théoriques de supposer que lorsqu'une masse subit une accélération (comme par exemple en se mouvant sur une orbite elliptique autour d'une autre masse, ou lors d'un effondrement gravitationnel), elle cède de l'énergie sous forme d'ondes gravitationnelles. Ces ondes, elles aussi, peuvent revêtir un aspect « particulier », et la particule gravitationnelle est appelée *graviton*.

La force gravitationnelle est considérablement plus faible que la force électromagnétique. La force d'attraction gravitationnelle mutuelle d'un proton et d'un électron ne dépasse pas $1/10^{39}$ fois leur force d'attraction électromagnétique. Le graviton a donc moins d'effets que le photon, et sa détection doit être d'une difficulté inimaginable.

Le physicien américain Joseph Weber s'est néanmoins attaqué en 1957 à la rude tâche de détecter le graviton. Il a utilisé une paire de cylindres en aluminium de 70 centimètres de diamètre et 1,5 mètre de longueur, chaque cylindre étant suspendu à un fil dans une chambre à vide. Les gravitons (qui devaient être détectés sous leur forme d'onde) devaient déplacer légèrement ces cylindres, et il a fallu faire appel à un système de mesure capable de détecter un déplacement de $1/10^{14}$ centimètre. Les faibles ondes des gravitons, venues de loin dans l'espace, doivent baigner toute notre planète et les deux cylindres, séparés par une grande distance, devaient être affectés simultanément. Weber a annoncé en 1969 qu'il avait détecté l'effet des ondes gravitationnelles. Il en est résulté une excitation intense dans le milieu des physiciens car cette observation confirmait une théorie particulièrement importante (la théorie de la relativité générale d'Einstein). Malheureusement, les contes scientifiques ne se terminent pas toujours bien. Aucun autre chercheur, quelle que soit la méthode essayée, n'est parvenu à reproduire les résultats de Weber, et l'opinion générale prévaut que les gravitons ne sont toujours pas détectés. Les physiciens ont néanmoins suffisamment confiance dans la théorie pour être convaincus de leur existence. Ce sont des particules de masse et charge nulles, de spin 2 ; ce sont donc aussi des bosons. Les gravitons ne figurent donc pas non plus dans la liste des leptons.

Les photons et les gravitons n'ont pas d'antiparticule associée, ou plutôt, chacun est sa propre antiparticule. On peut se représenter cela en imaginant une feuille de papier pliée dans le sens de la longueur, puis dépliée, de sorte qu'il reste un pli marqué selon l'axe de la feuille. Un petit cercle dessiné à gauche du pli et un autre cercle, à égale distance à droite du pli, représenteraient un électron et un positron. Le photon et le graviton seraient juste sur le pli.

---

* Il s'agit plus précisément de la conservation du moment angulaire total, résultant des moments cinétiques (dus aux mouvements orbitaux), qui ne prennent que des valeurs entières, et des spins des particules (moments cinétiques propres), qui peuvent prendre des valeurs entières ou demi-entières. (*N.D.T.*)

NEUTRINOS ET ANTINEUTRINOS

Il semblerait, jusqu'à présent, qu'il n'y ait que deux leptons, l'électron et le positron. Les physiciens se seraient volontiers contentés de cette situation ; il ne semblait y avoir aucun besoin pressant pour quoi que ce soit de plus... n'étaient les problèmes posés par l'émission de particules bêta par des noyaux radioactifs.

La particule émise par un noyau radioactif emporte généralement une quantité d'énergie appréciable. D'où vient cette énergie ? Elle résulte de la conversion d'une petite part de la masse du noyau ; en d'autres termes, en expulsant une particule, le noyau perd toujours un peu de masse. Mais les physiciens étaient depuis longtemps intrigués par le fait que la particule bêta émise lors de la désintégration d'un noyau n'emportait généralement pas assez d'énergie pour rendre compte de la masse perdue par le noyau. En fait, les électrons ne présentaient pas tous le même déficit. Ils émergeaient avec un large spectre d'énergies, dont le maximum (atteint par très peu d'électrons) avait presque la bonne valeur, mais il manquait peu ou prou quelque chose à toutes les autres. Et cet état de choses n'était pas sytématiquement associé à l'émission de particules subatomiques. Les particules alpha émises par un nuclide particulier avaient toutes la même valeur d'énergie attendue. Qu'est-ce qui n'allait pas dans l'émission bêta ? Qu'était-il arrivé à l'énergie manquante ?

C'est Lise Meitner qui fut la première, en 1922, à insister sur cette question, de façon pressante, et en 1930 Niels Bohr, entre autres, était prêt à abandonner le grand principe de la conservation de l'énergie, tout au moins en ce qui concernait les particules subatomiques. Mais en 1931, pour sauver la conservation de l'énergie (voir le chapitre 8), Wolfgang Pauli proposait une solution à l'énigme de l'énergie manquante. Son explication est des plus simples : une autre particule sort du noyau en même temps que la particule bêta, en emportant l'énergie manquante. Cette seconde particule mystérieuse a des propriétés plutôt étranges : sa charge et sa masse sont nulles ; tout ce qu'elle possède, tandis qu'elle file à la vitesse de la lumière, c'est une certaine quantité d'énergie. Cette particule avait tout l'air d'être un objet fictif inventé juste pour équilibrer le bilan d'énergie.

Et pourtant, on avait à peine émis l'hypothèse de cette particule que les physiciens étaient certains de son existence. Lorsque le neutron fut découvert et que l'on s'aperçut qu'il se désintégrait en un proton, en libérant un électron qui, comme dans la désintégration bêta, manquait d'énergie, ils en furent encore plus convaincus. Enrico Fermi, en Italie, donna à cette particule en puissance le nom de *neutrino*, qui signifie en italien le « petit neutre ».

Le neutron a apporté aux physiciens une nouvelle preuve de l'existence du neutrino. Comme je l'ai dit, presque toutes les particules ont un spin. La grandeur du spin s'exprime en multiples de 1/2, positifs ou négatifs selon le sens du spin. Le proton, le neutron et l'électron ont chacun un spin qui vaut 1/2. Dans ces conditions, si le neutron, avec son spin 1/2, donne naissance à un proton et un électron, chacun de spin 1/2, qu'advient-il de la loi de conservation du spin ? Il y a quelque chose qui ne va pas. Le proton et l'électron peuvent ajouter leurs spins pour donner la valeur 1 (si les deux spins ont même sens) ou 0 (si les spins sont opposés) ; mais de quelque façon que l'on s'y prenne, leurs spins ne peuvent donner la valeur 1/2. A nouveau, c'est le neutrino qui vient à la rescousse. Prenons

le cas d'un neutron de spin +1/2. Soit un proton de spin +1/2 et un électron de spin -1/2 ; le spin résultant vaut 0. Attribuons maintenant au neutrino le spin +1/2, ce qui en fait aussi un fermion (et donc un lepton), et le bilan est parfaitement équilibré :

+1/2 (n) = +1/2 (p) -1/2 (e) +1/2 (neutrino)

Mais il reste encore une opération d'équilibrage à réaliser. Une particule unique (le neutron) a formé deux particules (le proton et l'électron) et, si nous incluons le neutrino, en fait trois particules. Il semble plus raisonnable de supposer que le neutron est converti en deux particules et une antiparticule, soit une particule résultante. En d'autres termes, ce dont nous avons besoin pour l'équilibrage n'est pas un neutrino, mais un antineutrino.

Le neutrino lui-même apparaîtrait dans la conversion d'un proton en neutron. Les produits seraient alors un neutron (une particule), un positron (une antiparticule) et un neutrino (une particule). Le bilan est encore équilibré.

En d'autres termes, l'existence du neutrino et de l'antineutrino sauverait non seulement une, mais trois lois de conservation importantes : la conservation de l'énergie*, la conservation du spin, et la conservation du nombre particule-antiparticules. La préservation de ces lois est importante car elles semblent valides dans toutes sortes de réactions nucléaires qui n'impliquent ni électrons, ni positrons, et il serait donc particulièrement pratique que leur validité s'étende aux réactions où interviennent aussi ces particules.

Les conversions de proton en neutron les plus importantes interviennent dans les réactions nucléaires qui se produisent dans le Soleil et les autres étoiles. Les étoiles émettent donc des flux intenses de neutrinos, et on estime que 6 à 8 % de leur énergie rayonnée part sous cette forme. Mais ceci n'est vrai que pour les étoiles du même type que notre Soleil. Le physicien américain Hong Yee Chiu a suggéré, en 1961, que lorsque la température du centre de l'étoile s'élève, il se produit d'autres réactions génératrices de neutrinos. A mesure que l'étoile s'achemine, suivant le cours de son évolution, vers un cœur plus chaud (voir le chapitre 2), une proportion croissante de son énergie rayonnée est emportée par les neutrinos.

Cette notion est d'une importance cruciale. Le processus ordinaire de transfert de l'énergie par les photons est lent. Les photons interagissent avec la matière et ils ne se fraient un chemin du cœur du Soleil à sa surface qu'au prix d'innombrables absorptions et réémissions. De ce fait, bien que la température centrale du Soleil soit de 15 000 000 °C, sa surface n'est qu'à 6 000 °C. La substance du Soleil est un bon isolant thermique.

Les neutrinos, en revanche, n'interagissent pratiquement pas avec la matière. On a calculé qu'un neutrino moyen pouvait traverser une épaisseur de 100 années-lumière de plomb en n'ayant que 50 % de chances d'être absorbé. Ainsi, tous les neutrinos formés dans le cœur du Soleil le quittent immédiatement à la vitesse de la lumière, pour en atteindre, sans autre interaction, la surface en moins de trois secondes et s'échapper. (Tout neutrino qui se dirige vers nous nous traverse sans nous affecter, que ce

* Ainsi que la conservation de la quantité de mouvement (ou impulsion), qui en est une conséquence nécessaire dans le cadre de la théorie de la relativité. (*N.D.T.*)

soit de jour ou de nuit, car la nuit, lorsque toute la Terre s'interpose entre le Soleil et nous, les neutrinos traversent la Terre puis notre corps aussi facilement que notre corps seul.)

Chiu a calculé que lorsque la température centrale atteint 6 000 000 000 °K, la plus grande partie de l'énergie est rayonnée par l'étoile sous forme de neutrinos. Les neutrinos s'échappent immédiatement, emportant leur énergie, et le centre de l'étoile se refroidit brutalement. C'est peut-être ce mécanisme qui conduit à la contraction catastrophique qui se manifeste sous forme de supernova.

À L'AFFÛT DU NEUTRINO

Les antineutrons sont produits par la conversion des neutrons en protons mais, pour autant que l'on sache, celle-ci ne se fait pas à un taux comparable à celui des réactions qui émettent les flux de neutrinos des étoiles. Les sources d'antineutrinos les plus importantes sont la radioactivité naturelle et la fission de l'uranium (que je traiterai plus en détail au chapitre 10).

Naturellement, les physiciens ne pouvaient pas en rester là : il leur fallait effectivement capturer le neutrino ; les scientifiques ne se contentent jamais d'accepter des phénomènes ou des lois de la nature entièrement par des actes de foi. Mais comment détecter une entité aussi évanescente que le neutrino, un objet sans masse, sans charge, et pratiquement sans propension à interagir avec la matière ordinaire ?

Pourtant, il y avait un léger espoir. Bien que la probabilité de réaction d'un neutrino avec une particule quelconque soit minuscule, elle n'est pas tout à fait nulle. Traverser sans coup férir 100 années-lumière de plomb donne une idée de la moyenne, mais il y a des neutrinos qui vont réagir avec une particule avant d'aller aussi loin, et il y a même une infime proportion du nombre total qui sera arrêtée par quelques millimètres de plomb, ou l'équivalent.

Un groupe de physiciens du laboratoire de Los Alamos, animé par Clyde Lorrain Cowan et Frederick Reines, a tenté l'impossible en 1953. Ils ont placé leur appareil de détection des neutrinos près d'un grand réacteur à fission de la Commission pour l'énergie atomique, sur les rives de la Savannah, dans l'Etat de Géorgie. Un réacteur produit des quantités de neutrons, qui doivent donc donner naissance à un flux d'antineutrinos. Pour les capturer, les expérimentateurs ont utilisé de grands réservoirs d'eau. L'idée était de bombarder les protons (les noyaux d'hydrogène) de l'eau avec les antineutrinos et de détecter les produits de capture d'un antineutrino par un proton.

A quoi fallait-il s'attendre ? Lorsqu'un neutron se désintègre, il donne un proton, un électron et un antineutrino. L'absorption d'un antineutrino par un proton devrait conduire à l'inverse. C'est-à-dire que le proton devrait être converti en neutron, en émettant un positron au cours du processus. Il y avait donc deux choses à chercher : la création de neutrons et la création de positrons. Les neutrons pouvaient être détectés dans une solution d'un des composés du cadmium dans l'eau, car en absorbant les neutrons le cadmium émet des rayons gamma d'énergie caractéristique. Et on pouvait identifier les positrons par leur interaction d'annihilation avec les électrons, qui donnait d'autres rayons gamma caractéristiques.

Avec des instruments sensibles aux rayons gamma ayant exactement ces deux énergies et séparés par l'intervalle de temps convenable, on pouvait être certain d'avoir attrapé des antineutrinos.

Les expérimentateurs disposèrent leurs astucieux détecteurs et attendirent patiemment ; en 1956, exactement un quart de siècle après l'invention de cette particule par Pauli, ils capturaient enfin l'antineutrino. Les journaux et même quelques publications savantes l'ont appelé simplement le *neutrino*.

Pour obtenir l'authentique neutrino, il nous faut une source de neutrinos intense. La plus évidente est le Soleil. Quel système peut-on utiliser pour détecter le neutrino, par opposition à l'antineutrino ? Un procédé (suivant une proposition du physicien soviétique d'origine italienne Bruno Pontecorvo) part du chlore 37, qui constitue environ un quart des atomes du chlore naturel. Son noyau contient 17 protons et 20 neutrons. Si l'un de ces neutrons absorbe un neutrino, il devient un proton (et émet un électron). Le noyau aura alors 18 protons et 19 neutrons, et sera de l'argon 37.

Pour constituer une cible de neutrons du chlore efficace, on peut utiliser le chlore liquide, mais c'est une substance très corrosive et toxique, dont le maintien à l'état liquide présente un problème de réfrigération. On peut aussi utiliser des composés organiques contenant du chlore ; en particulier, le tétrachloroéthylène convient.

Le physicien américain Raymond R. Davis a utilisé ce piège à neutrinos en 1956 pour prouver que le neutrino et l'antineutrino sont bien deux choses différentes. En supposant ces deux particules différentes, le piège ne devait détecter que les neutrinos et pas les antineurinos. Placé près d'un réacteur à fission, en 1956, dans des conditions où il y avait certainement des antineutrinos, il n'en a détecté aucun.

Il ne restait plus qu'à tenter de détecter les neutrinos du Soleil. Un énorme réservoir de 400 000 litres de tétrachloroéthylène fut utilisé pour cela. Le réservoir était placé au fond d'une mine dans le Dakota du Sud. Il était donc protégé par une épaisseur suffisante pour que toute particule venant du Soleil soit absorbée, sauf les neutrinos. (Nous avons ici un exemple de situation paradoxale où, pour étudier le Soleil, il faut s'enfouir profondément dans les entrailles de la Terre.) Pour permettre l'accumulation d'une quantité détectable d'argon 37, le réservoir resta plusieurs mois exposé aux neutrinos solaires. Il fut ensuite rincé à l'hélium pendant vingt-deux heures ; il n'y avait plus qu'à mesurer la petite quantité d'argon 37 contenue dans l'hélium gazeux. Des neutrinos solaires furent ainsi détectés en 1968, mais seulement un tiers du nombre prédit par les théories sur ce qui se passait à l'intérieur du Soleil. C'est là un résultat bien gênant sur lequel je reviendrai plus tard dans ce chapitre.

L'INTERACTION NUCLÉAIRE

Notre liste des particules subatomiques en comprend maintenant dix : quatre particules massives (ou *baryons*, d'après un mot grec qui signifie « lourd ») – le proton, le neutron, l'antiproton et l'antineutron ; quatre leptons – l'électron, le positron, le neutrino et l'antineutrino ; deux bosons – le photon et le graviton. Et pourtant, ce n'est pas assez, comme les physiciens en ont décidé après les considérations suivantes.

On peut facilement expliquer par l'*interaction électromagnétique* les forces ordinaires d'attraction entre un proton et un électron isolés, ou de répulsion entre deux protons ou deux électrons. Le mode de liaison de deux atomes ou de deux molécules s'explique aussi par l'interaction électromagnétique, l'attraction entre noyaux chargés positivement et électrons extérieurs.

Tant que l'on croyait le noyau atomique constitué de protons et d'électrons, on était en droit d'estimer que l'interaction électromagnétique, l'attraction nette entre protons et électrons, suffirait à expliquer pourquoi le noyau était aussi solide. Mais lorsqu'il fallut accepter la théorie proton-neutron pour la structure nucléaire, en 1930, les physiciens furent consternés de réaliser que la cohésion des éléments du noyau n'avait aucune explication.

Les seules particules chargées présentes dans le noyau étant des protons, l'interaction électromagnétique devrait se traduire par de violentes répulsions entre les protons étroitement serrés dans le minuscule noyau. Dès l'instant qu'il est formé (si tant est que l'on puisse même le former), tout noyau chimique devrait tomber en pièces avec une violence explosive.

A l'évidence, un autre type d'interaction doit intervenir ; quelque chose de beaucoup plus fort que l'interaction électromagnétique et capable de la dominer. En 1930, la seule autre interaction connue était l'*interaction gravitationnelle*, tellement plus faible que l'électromagnétique que l'on peut pratiquement la négliger dans les processus subatomiques. Il devait donc exister une sorte d'*interaction nucléaire*, jusqu'alors inconnue mais extrêmement puissante.

Cette intensité de l'interaction nucléaire est prouvée par la considération suivante. Dans un atome d'hélium, on peut séparer les deux électrons du noyau en dépensant une énergie de 54 électron-volts, quantité d'énergie qui suffit pour caractériser une des manifestations les plus rigoureuses de l'interaction électromagnétique.

D'un autre côté, pour séparer le proton et le neutron d'un deuton, l'un des noyaux les moins liés, il faut deux millions d'électron-volts. Même en tenant compte du fait que les particules dans le noyau sont beaucoup plus proches les unes des autres que dans l'atome, on peut estimer sans risque que l'interaction nucléaire est environ 130 fois plus forte que l'interaction électromagnétique.

Mais quelle est la nature de cette interaction nucléaire ? La première idée féconde est venue de Werner Heisenberg en 1932, lorsqu'il a suggéré que les protons étaient maintenus ensemble par des *forces d'échange*. Dans son modèle, les protons et les neutrons du noyau échangent continuellement leur identité, de sorte qu'une particule donnée est d'abord un proton, puis un neutron, et ainsi de suite. Ce mécanisme peut assurer la stabilité du noyau de la même façon que l'on tient une pomme de terre chaude en la faisant passer rapidement d'une main à l'autre. Avant qu'un proton ait pu « réaliser » (pour ainsi dire) qu'il est un proton et tenter de s'éloigner de ses voisins protons, il est devenu neutron et peut rester là où il est. Naturellement, le proton ne peut s'en tirer comme cela que si l'échange se déroule extrêmement rapidement, disons en 10^{-18} seconde.

Une autre façon d'envisager cette interaction consiste à imaginer que deux particules en échangent une troisième. Chaque fois que la particule A émet la particule d'échange, elle recule pour conserver l'impulsion totale. Chaque fois que la particule B absorbe la particule échangée, elle recule pour la même raison. Tandis que la particule d'échange rebondit de l'une

à l'autre, les particules A et B s'éloignent de plus en plus, si bien qu'elles semblent subir une sorte de répulsion. Si, en revanche, la particule échangée se déplace comme un boomerang, de l'arrière de la particule A vers l'arrière de la particule B, alors les deux particules se rapprocheront et sembleront soumises à une attraction.

Selon la théorie de Heisenberg, il semble que toutes les forces d'attraction ou de répulsion résultent d'échanges de particules. Dans le cas de l'attraction et de la répulsion électromagnétiques, la particule échangée est le photon ; pour l'attraction gravitationnelle (il n'existe apparemment pas de répulsion gravitationnelle), la particule échangée est le graviton.

Le photon et le graviton ont une masse nulle, et c'est pour cette raison que les forces électromagnétique et gravitationnelle ne décroissent que selon l'inverse du carré de la distance et qu'elles peuvent donc s'exercer sur des espaces considérables.

L'interaction gravitationnelle et l'interaction électromagnétique sont des *interactions à longue portée* et, pour autant que l'on sache à ce jour, les seules de ce type.

Dès lors que l'on supposait son existence, l'interaction nucléaire ne pouvait appartenir à ce type. Pour que le noyau subsiste, il fallait qu'elle soit très intense à l'intérieur du noyau ; mais elle était pratiquement indétectable hors du noyau, sinon elle aurait été découverte depuis longtemps. L'intensité de l'interaction nucléaire devait donc décroître rapidement avec la distance. Pour tout doublement de la distance, il lui fallait peut-être se trouver divisée par 100, au lieu de 4 comme dans le cas des interactions électromagnétique et gravitationnelle. Pour cette raison, aucune particule de masse nulle ne pouvait faire l'affaire.

LE MUON

En 1935, le physicien japonais Hideki Yukawa a analysé mathématiquement le problème. Une particule d'échange douée de masse créerait un champ de force à courte portée. La masse serait inversement proportionnelle à la portée : plus grande la masse, plus courte la portée. La masse convenable devait être quelque part entre celle du proton et celle de l'électron ; Yukawa l'a estimée entre 200 et 300 fois la masse de l'électron.

A peine un an s'était écoulé que l'on découvrait précisément cette particule. Carl Anderson (qui avait découvert le positron), étudiant les traces laissées par des rayons cosmiques secondaires, à l'Institut de technologie de Californie, remarqua une courte trace plus courbée que celle d'un proton et moins courbée que celle d'un électron. En d'autres termes, la particule ainsi révélée avait une masse intermédiaire. Bientôt on détecta d'autres traces, et ces particules furent baptisées *mésotrons*, ou *mésons* en abrégé.

On a découvert plus tard d'autres particules dans ce domaine des masses intermédiaires, et on distingue cette première particule sous le nom de *méson mu* ou *muon* (« mu » est une des lettres de l'alphabet grec ; elles ont maintenant presque toutes été utilisées pour désigner des particules subatomiques). Comme dans le cas des particules déjà rencontrées, le muon existe sous deux formes, positive et négative.

Le muon négatif, qui a 206,77 fois la masse de l'électron (et donc un neuvième de la masse du proton) est une particule ; le muon positif est

l'antiparticule. Les muons négatif et positif correspondent respectivement à l'électron et au positron. En 1960, il était en fait devenu évident que le muon négatif était identique en tout point – excepté sa masse – à l'électron. C'était un *électron lourd*. De même, le muon positif était un positron lourd.

Les muons positif et négatif peuvent s'annihiler mutuellement et, avant cela, tourner temporairement autour l'un de l'autre, tout comme des électrons positif et négatif. Le physicien américain Vernon Willard Hughes a découvert en 1960 une variante de cette situation. Il a détecté un système où l'électron tourne autour d'un muon positif, un système qu'il a appelé *muonium*. (Un positron en orbite autour d'un muon négatif constituerait un *antimuonium*.)

L'atome de muonium (si l'on peut l'appeler ainsi) est tout à fait analogue à l'hydrogène 1, dans lequel un électron est en orbite autour d'un proton positif, et leurs propriétés présentent de nombreuses similitudes. En dépit de l'identité apparente du muon et de l'électron, leur différence de masse suffit pour qu'ils ne soient pas rigoureusement symétriques l'un de l'autre, et ils ne peuvent s'annihiler mutuellement. Le muonium n'est donc pas affecté de la même instabilité que le positronium. Le muonium dure longtemps et, laissé en paix, durerait éternellement, si ce n'était le fait que le muon lui-même a une fin car, comme je vais bientôt en parler, il est très instable.

Il y a une autre similitude entre muons et électrons : de la même façon que les particules lourdes peuvent produire des électrons et des antineutrinos (par exemple lorsqu'un neutron est converti en proton), elles peuvent interagir pour donner des muons négatifs et des antineutrinos ou des muons positifs et des neutrinos. Pendant des années, les physiciens n'ont jamais douté que les neutrinos qui accompagnent les muons et ceux qui accompagnent les électrons fussent identiques. Mais en 1962 on s'est aperçu que les neutrinos ne sont pas interchangeables ; le neutrino de l'électron n'intervient jamais dans une réaction où se forme un muon, et le neutrino du muon ne joue aucun rôle dans les interactions où apparaît un électron ou un positron.

En bref, les physiciens se retrouvaient à la tête de deux paires de particules sans charge ni masse, l'antineutrino de l'électron et le neutrino du positron, plus l'antineutrino du muon négatif et le neutrino du muon positif. On est actuellement incapable d'exprimer la différence entre les deux neutrinos et entre les deux antineutrinos, mais ils sont différents.

Les muons diffèrent de l'électron et du positron sur un autre point, celui de la stabilité. Abandonnés à eux-mêmes, l'électron ou le positron restent indéfiniment inchangés. En revanche, le muon est instable et se désintègre après une durée de vie moyenne de quelques millionièmes de seconde. Le muon négatif finit en électron (plus un antineutrino de type électronique et un neutrino de type muonique), tandis que le muon positif fait exactement l'opposé en produisant un positron, un neutrino électronique et un antineutrino muonique.

Lorsqu'un muon se désintègre il forme donc un électron, 200 fois plus léger, et deux neutrinos sans aucune masse. Qu'advient-il des 99,9 % restants de la masse ? Ils sont évidemment convertis en énergie cinétique des particules sortantes.

Inversement, en concentrant suffisamment d'énergie dans un minuscule volume, on peut assister à la formation non pas d'une paire électron-

positron, mais d'une paire plus coûteuse, une paire semblable à la paire électron-positron, si ce n'est son excédent d'énergie qui se manifeste sous forme de masse. Cet excès de masse n'est pas très lié à l'électron et au positron de base, de sorte que le muon est instable et s'en débarrasse rapidement pour devenir un électron ou un positron.

LE TAU

Naturellement, en concentrant encore plus d'énergie dans un petit volume, on obtient un électron encore plus massif. Martin L. Perl a détecté la preuve de l'existence de cet électron superlourd en 1974, au moyen d'un accélérateur qui réalisait des collisions frontales d'électrons et de positrons de haute énergie, en Californie. Il a appelé cette particule *l'électron tau,* que l'on abrège généralement en *tau* (c'est encore une autre lettre de l'alphabet grec).

Comme on pouvait s'y attendre, le tau a environ 17 fois la masse du muon, et donc environ 3 500 fois la masse d'un électron. En fait, le tau est deux fois plus lourd qu'un proton ou un neutron. En dépit de sa masse,le tau est un lepton car, à part cette masse et son instabilité, il a toutes les propriétés d'un électron. Avec toute cette masse, on pouvait s'attendre à ce qu'il soit beaucoup plus instable que le muon, et c'est bien le cas. Le tau ne vit qu'en moyenne un milliardième de seconde avant de se désintégrer en muon, puis en électron.

Il existe bien sûr un tau négatif et un tau positif, et les physiciens sont convaincus de l'existence d'un troisième type de neutrino et d'antineutrino, même si ceux-ci n'ont pas encore été explicitement détectés.

LA MASSE DU NEUTRINO

On connaît maintenant douze leptons : les électrons négatif et positif (ce dernier étant appelé positron), les muons négatif et positif, les taus négatif et positif, le neutrino et l'antineutrino électroniques, le neutrino et l'antineutrino muoniques et le neutrino et l'antineutrino associés au tau. Ils sont divisés en trois catégories (ou, comme disent maintenant les physiciens, trois *saveurs**). Il y a l'électron et son neutrino associé et leurs antiparticules, le muon et son neutrino associé et leurs antiparticules, le tau et son neutrino associé et leurs antiparticules.

Connaissant déjà trois saveurs, il n'y a guère de raison pour qu'il n'en existe pas d'autres. Il se peut qu'en augmentant indéfiniment l'énergie, on puisse former de plus en plus de saveurs de leptons, chaque saveur étant plus massive et moins stable que la précédente. Bien qu'il n'y ait aucune limite théorique au nombre de saveurs, il y aurait bien sûr une limite pratique. Il faudrait peut-être à la fin toute l'énergie de l'Univers pour former un lepton particulièrement élevé dans la hiérarchie, et on ne pourrait aller plus loin ; cette particule serait tellement instable que son existence n'aurait aucune signification.

---

* Le terme de « saveur » est rarement utilisé par les spécialistes à propos des leptons, et ils distinguent alors six saveurs en tout (comme pour les quarks, voir plus loin) pour caractériser les trois leptons massifs et les trois neutrinos. (*N.D.T.*)

Si nous en restons aux trois saveurs actuellement connues, le mystère des neutrinos s'épaissit. Comment peut-il y avoir trois paires de fermions de charge et de masse nulles, ayant chacun des interactions essentiellement différentes avec les autres particules alors qu'on ne leur trouve aucune caractéristique différente ?

Peut-être existe-t-il une propriété distinctive que nous n'avons pas convenablement cherchée ? Nous admettons par exemple que les trois saveurs de neutrino ont une masse nulle et se déplacent donc toujours à la vitesse de la lumière. Supposons plutôt que chaque saveur de neutrino ait une minuscule masse, différente des masses des deux autres. Dans ce cas, les propriétés seraient légèrement modifiées d'un neutrino à l'autre. Par exemple, ils iraient chacun à une vitesse un peu inférieure à la vitesse de la lumière, l'écart par rapport à la vitesse de la lumière étant légèrement différent pour chacun d'eux.

On a certaines raisons de supposer que, dans ce cas, tout neutrino change d'identité en cours de route, pouvant être un neutrino électronique à un moment donné, un neutrino muonique un peu plus tard et un neutrino taunique encore à un autre moment. Ces changements constituent ce que l'on appelle les *oscillations de neutrino*, initialement proposées par un groupe de physiciens japonais en 1963.

Vers la fin des années soixante-dix, Frederick Reines, l'un des premiers à avoir détecté les neutrinos, a entrepris de tester cette hypothèse en compagnie de Henri W. Sobel et Elaine Pasierb, de l'université de Californie. Ils ont utilisé environ 300 kilos d'eau lourde très pure qu'ils ont bombardée avec des neutrinos venant de la fission de l'uranium. Ce mécanisme ne devait produire *que* des neutrinos électroniques.

Les neutrinos peuvent induire deux sortes d'événements. Un neutrino peut frapper le couple neutron-proton du noyau d'hydrogène lourd dans l'eau lourde, les séparer, et continuer son chemin. C'est une *réaction de courant neutre*, que n'importe quelle saveur de neutrino peut déclencher. D'autre part, le neutrino peut, en tapant dans la combinaison neutron-proton, induire une transformation du proton en neutron, produisant un électron, et dans ce cas le neutrino disparaît. C'est une *réaction de courant chargé* que *seuls* des neutrinos électroniques peuvent déclencher.

On peut calculer le nombre d'événements de chaque type qui se produiront selon que les neutrinos n'oscillent pas et restent des neutrinos électroniques, ou qu'ils oscillent et qu'un certain nombre changent de nature. Reines a annoncé en 1980 que son expérience semblait indiquer l'existence d'oscillations de neutrinos. (Je dis bien « semblait », car l'effet était à la limite du mesurable, et les expérimentateurs qui ont vérifié la chose déclarent n'avoir détecté *aucun* signe d'oscillation.)*

La question reste douteuse, mais des expériences réalisées par des physiciens de Moscou, par un procédé qui n'a rien à voir avec les oscillations, semblent indiquer que le neutrino électronique pourrait avoir une masse de 40 électron-volts. Sa masse ne serait alors que de 1/13 000 de la masse de l'électron, et il n'y aurait rien d'étonnant à ce que cette particule ait pu passer pour dénuée de toute masse. Si Reines a raison

---

* En 1984, un groupe de physiciens d'Annecy et de Grenoble a obtenu, avec les antineutrinos fournis par l'E.D.F. dans son réacteur du Bugey, des résultats indiquant des oscillations de neutrinos. (*N.D.T.*)

et s'il y a bien une oscillation de neutrino, celle-ci expliquerait la pénurie de neutrinos d'origine solaire que j'ai déjà mentionnée dans ce chapitre et qui intrigue tant les scientifiques. Le système utilisé par Davis pour détecter les neutrinos solaires n'est sensible qu'aux neutrinos électroniques. Si les neutrinos émis par le Soleil oscillent de sorte qu'ils arrivent sur Terre à l'état de mélange des trois saveurs en proportions du même ordre, il est normal de ne détecter qu'un tiers des neutrinos attendus.

De plus, si les neutrinos ont une petite masse, même si ce n'est que 1/13 000 de la masse d'un électron, le calcul montre qu'ils abondent tellement dans l'Univers qu'ils dominent de beaucoup l'ensemble des protons et des neutrons. Plus de 99 % de la masse de l'Univers viendrait des neutrinos, et ils pourraient facilement rendre compte de la « masse manquante » dont je parle dans le chapitre 2. Il y aurait en fait une masse de neutrinos dans l'Univers suffisante pour fermer celui-ci et nous assurer que finalement l'expansion cessera et que l'Univers se contractera à nouveau.

Tout cela *si* Reines a raison. Mais nous ne savons pas encore.

Hadrons et quarks

Comme le muon est finalement une sorte d'électron lourd, il ne peut guère jouer le rôle du liant nucléaire que cherchait Yukawa. On ne trouve pas d'électrons dans le noyau, il ne devrait donc pas y avoir plus de muons. On savait cela, de façon purement empirique, bien avant d'avoir suspecté la parenté du muon et de l'électron ; tout simplement, les muons ne montraient aucune tendance à interagir avec les noyaux. Pendant un certain temps, la théorie de Yukawa a semblé vaciller.

LES PIONS ET LES MÉSONS

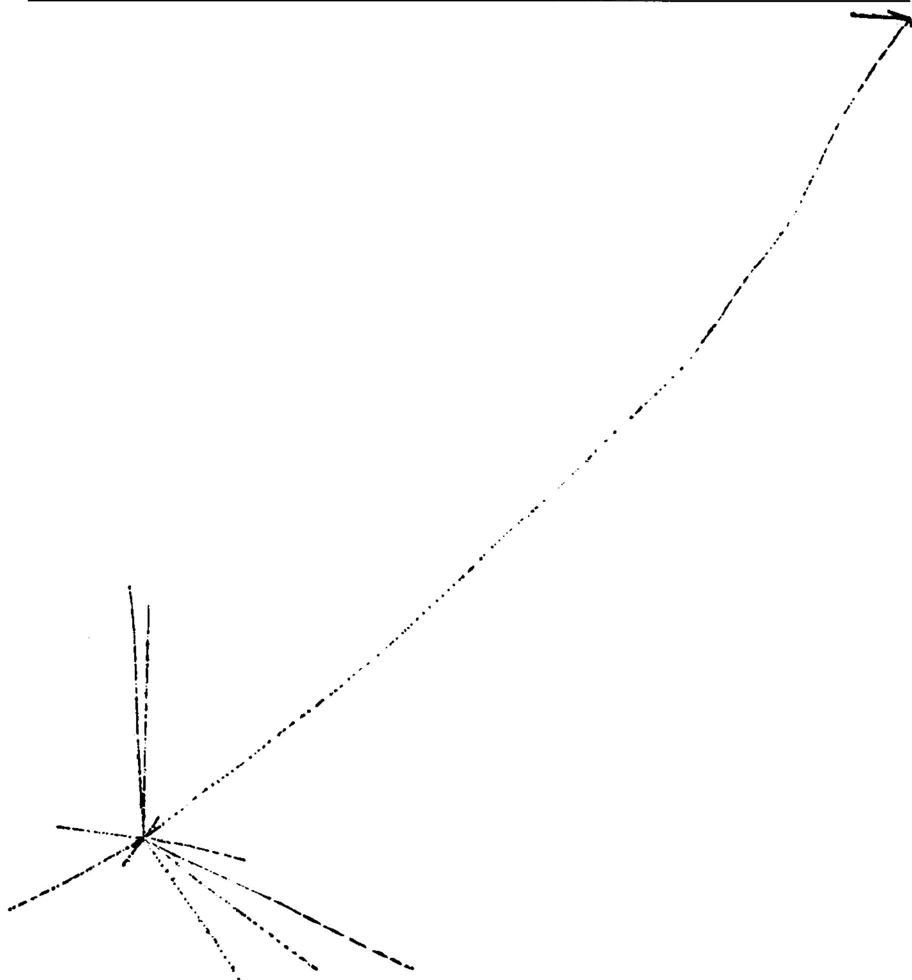
Mais en 1947, le physicien britannique Cecil Frank Powell a découvert un autre type de méson dans les photographies de rayons cosmiques. Celui-ci était un peu plus massif que le muon et s'est avéré avoir 273 fois la masse de l'électron. Le nouveau méson a été appelé *méson pi* ou *pion.*

On a trouvé que le pion réagissait fortement avec le noyau ; c'était donc bien la particule annoncée par Yukawa. (Le prix Nobel de physique est venu récompenser Yukawa en 1949 et Powell en 1950.) Comme prévu, il y avait un pion positif qui jouait le rôle de force d'échange entre protons et neutrons, et l'antiparticule correspondante, le pion négatif, qui remplissait la même tâche vis-à-vis des antiprotons et antineutrons. Ils sont tous deux encore plus éphémères que les muons ; après une vie moyenne d'un quarantième de microseconde, le pion se désintègre en muon et neutrino muonique. (Et, bien sûr, le muon se désintègre ensuite en électron et en neutrinos supplémentaires.) Il existe aussi un pion neutre qui est sa propre antiparticule. (En d'autres termes, il n'y a qu'une seule variété de cette particule.) Il est extrêmement instable et se désintègre après moins de 10^{-18} seconde en une paire de rayons gamma.

En dépit de son caractère nucléaire, le pion peut tourner autour d'un noyau avant d'interagir avec, pour former parfois un *atome pionique* que l'on a détecté en 1952. On peut en fait faire tourner l'une autour de l'autre n'importe quelle paire de particules, ou de systèmes de particules, positif et négatif ; au cours des années soixante, les physiciens ont ainsi étudié un certain nombre d'« atomes exotiques » éphémères pour glaner des informations sur la structure des particules.

Les pions n'étaient que les premiers découverts de toute une classe de particules regroupées sous le terme de *mésons*. Le muon ne figure plus parmi les mésons, bien qu'il ait été la première particule à laquelle on ait attribué ce nom. Les mésons interagissent fortement avec les protons et les neutrons, contrairement aux muons qui ont ainsi perdu tout droit d'appartenir au club.

Figure 7.8. Collision d'un méson et d'un noyau. Un méson de haute énergie, appartenant au rayonnement cosmique secondaire, a frappé un noyau et produit une étoile de mésons et de particules alpha *(en bas à gauche)* ; le méson de haute énergie a décrit ensuite le chemin hésitant vers le coin supérieur droit où une collision avec un autre noyau l'a finalement arrêté.



Comme particules membres du groupe autres que les pions, nous avons par exemple les *mésons K* ou *Kaons.* Ils ont été détectés pour la première fois en 1952 par deux physiciens polonais, Marian Danysz et Jerzy Pniewski. Ces particules ont 970 fois la masse de l'électron et donc à peu près la moitié de la masse d'un proton ou d'un neutron. Il existe deux variétés de méson K, le kaon positif et le kaon neutre, chacun avec son antiparticule associée. Naturellement, ils sont instables et se désintègrent en pion après environ une microseconde.

LES BARYONS

Au-dessus des mésons, il y a les baryons (un terme que j'ai déjà mentionné), qui comprennent le proton et le neutron. C'étaient les seuls spécimens connus dans le début des années cinquante. Mais on a découvert à partir de 1954 une série de particules encore plus massives (parfois appelées *hypérons*). Ce sont surtout les baryons qui ont proliféré ces dernières années, et le proton et le neutron ne sont que les plus légers des membres d'une grande famille.

Les physiciens ont découvert qu'il existe une loi de *conservation du nombre baryonique,* car dans toutes les désintégrations de particules, le nombre net de baryons (c'est-à-dire le nombre de baryons moins le nombre d'antibaryons) reste constant. La désintégration se produit toujours d'une particule lourde vers une particule moins lourde, ce qui explique que le proton soit stable et que ce soit le *seul* baryon stable. Il se trouve être le plus léger des baryons. S'il se désintégrait, il lui faudrait cesser d'être un baryon, ce qui le mettrait en infraction avec la loi de conservation du nombre baryonique. L'antiproton est stable pour la même raison, car c'est l'antibaryon le plus léger. Un proton et un antiproton peuvent évidemment s'unir dans une mutuelle annihilation car, pris ensemble, cela fait un baryon plus un antibaryon et un nombre baryonique total nul.

(Il existe aussi une *loi de conservation du nombre leptonique* qui rend compte du fait que l'électron et le positron sont les *seuls* leptons stables. Ce sont les leptons les moins lourds et ils ne peuvent se désintégrer en quoi que ce soit de plus léger sans violer cette loi de conservation. Il y a en fait une autre raison qui empêche l'électron et le positron de se désintégrer : ce sont les particules les moins lourdes qui possèdent une charge électrique. S'ils se désintègrent en quelque chose de plus simple, ils perdent cette charge électrique, disparition prohibée par la *loi de conservation de la charge électrique*. Il s'agit là, à vrai dire, d'une loi de conservation plus stricte que la conservation du nombre baryonique, comme nous le verrons, de sorte qu'électrons et positrons sont, d'une certaine façon, plus stables que les protons et antiprotons ou, au moins, ils *peuvent* être plus stables.)

On a donné des noms grecs aux premiers baryons découverts après le proton et le neutron. Il y a eu la *particule lambda*, la *particule sigma* et la *particule xi.* De la première il n'existe qu'une variété, une particule neutre ; de la seconde il y a trois variétés, positive, négative et neutre ; la troisième existe sous deux variétés, négative et neutre. Chacune de celles-ci a son antiparticule associée, ce qui nous fait en tout une douzaine de particules. Elles sont toutes très instables ; aucune ne vit plus d'un milliardième de seconde, et la particule sigma neutre se désintègre après 10^{-9} seconde.

La particule lambda étant neutre, elle peut remplacer un neutron dans un noyau pour former un *hypernoyau,* un système qui dure moins de 10^{-9} seconde. Le premier hypernoyau découvert était un noyau d'hypertritium, constitué d'un proton, d'un neutron et d'un lambda. Il a été trouvé par Danysz et Pniewski en 1952 parmi les produits du rayonnement cosmique. En 1963, Danysz a annoncé la découverte d'un hypernoyau contenant deux particules lambda. De plus, on peut substituer des hypérons négatifs aux électrons de l'édifice atomique, ce qui a été réalisé en 1968. Les orbites de ces succédanés massifs des électrons rasent le noyau de si près qu'ils passent en fait leur temps dans ses régions superficielles.

Mais toutes ces particules sont relativement stables ; elles vivent assez longtemps pour être directement détectables et se voir facilement attribuer une durée de vie et une personnalité en propre. Dans les années soixante, la première de toute une série de particules a été détectée par Alvarez (ce qui lui a valu le prix Nobel de physique en 1968). Ces particules ont une vie si brève que la seule preuve de leur existence réside dans la nécessité de rendre compte de leurs produits de désintégration. Leur durée de vie est de l'ordre de 10^{-24} seconde, et on peut se demander s'il s'agit là réellement de particules individualisées ou simplement de combinaisons de deux ou plusieurs particules qui marquent un temps d'arrêt pour se saluer avant de passer leur chemin.

Ces entités à très courte durée de vie sont appelées des *particules résonnantes* et, à mesure que les physiciens disposaient d'énergies de plus en plus élevées, on n'a cessé d'en trouver de nouvelles, si bien que l'on en connaît maintenant environ 150. Elles se rangent toutes parmi les mésons ou les baryons, et on regroupe les deux familles sous le nom d'*hadrons* (d'après un mot grec qui signifie « massif »). Les leptons, eux, s'en tiennent à trois saveurs, chaque saveur comprenant particule, antiparticule, neutrino et antineutrino.

Les physiciens devenaient de plus en plus mécontents devant cette profusion d'hadrons, comme les chimistes l'avaient été devant la multiplicité des éléments un siècle plus tôt. On commençait à avoir l'impression que les hadrons devaient être constitués de particules plus simples. A la différence des leptons, les hadrons ne sont pas des points ; ils ont un diamètre pas très gros, de l'ordre de 10^{-15} mètre, mais cela suffit pour que ce ne soient pas des points.

Le physicien américain Robert Hofstadter, dans les années cinquante, a étudié les noyaux avec des électrons de très haute énergie. Les électrons ricochaient plutôt que d'interagir avec les noyaux et, d'après le rebond, Hofstadter est parvenu à des conclusions sur la structure des hadrons finalement inadéquates, mais qui constituaient néanmoins un bon point de départ. Ce travail lui a valu d'être l'un des lauréats du prix Nobel de physique en 1961.

LA THÉORIE DES QUARKS

Une sorte de tableau périodique pour les particules subatomiques semblait nécessaire, un tableau où elles seraient regroupées en familles constituées par des membres de base et par d'autres particules qui seraient des états excités de ces membres de base (tableau 7.1).

Tableau 7.1.

Les particules subatomiques de longue durée de vie.

Nom de famille	*Nom de la particule*	*Symbole*	*Masse*	*Spin*	*Charge électrique*	*Antiparticule*	*Nombre de particules distinctes*	*Durée de vie moyenne (en secondes)*	*Mode de désintégration principal*
	Photon	γ	0	1	neutre	lui-même	1	infinie	—
	Graviton	—	0	2	neutre	lui-même	1	infinie	—
Électrons	Neutrino électronique	ν_e	0	½	neutre	$\overline{\nu_e}$	2	infinie	—
	Électron	e^-	1	½	négative	e^+ (positron)	2	infinie	—
Muons	Neutrino muonique	ν_μ	0(?)	½	neutre	$\overline{\nu_\mu}$	2	infinie	—
	Muon	μ^-	206,77	½	négative	μ^+	2	$2{,}197 \times 10^{-6}$	$e^- + \overline{\nu_e} + \nu_\mu$
Mésons	Pion	π^+	273,2	0	positive	π^-		$2{,}60 \times 10^{-8}$	$\mu^+ + \nu_\mu$
		π^-	273,2	0	négative	π^+	3	$2{,}60 \times 10^{-8}$	$\mu^- + \overline{\nu_\mu}$
		π^0	264,2	0	neutre	lui-même		$0{,}83 \times 10^{-16}$	$\gamma + \gamma$
	Kaon	K^+	966,6	0	positive	$\overline{K^+}$ (négatif)	4	$1{,}24 \times 10^{-8}$	$\mu^+ \; \nu_\mu$
		K^0	974	0	neutre	$\overline{K^0}$		$0{,}89 \times 10^{-10}$	$\pi^+ + \pi^-$
								et $5{,}18 \times 10^{-8}$	$\pi^+ + \pi^- + \pi^0$
Baryons	Nucléon	p (proton)	1836,12	½	positive	$\overline{p}$ (négative)	4	$> 10^{32}$ années	
		n (neutron)	1838,65	½	neutre	$\overline{n}$		898	$p + e^- + \overline{\nu_e}$
	Lambda	Λ^0	2182,8	½	neutre	$\overline{\Lambda^0}$	2	$2{,}63 \times 10^{-10}$	$P + \pi^-$
	Sigma	Σ^+	2327,7	½	positive	$\overline{\Sigma^+}$ (négatif)		$0{,}8 \times 10^{-10}$	$p + \pi^0$
		Σ^-	2340,5	½	négative	$\overline{\Sigma^-}$ (positif)	6	$1{,}48 \times 10^{-10}$	$n + \pi^-$
		Σ^0	2332	½	neutre	$\overline{\Sigma^0}$		6×10^{-20}	$\Lambda^0 + \gamma$
	Xi	Ξ^-	2580	½	négative	$\overline{\Xi^-}$ (positif)		$1{,}64 \times 10^{-10}$	$\Lambda^0 + \pi^-$
		Ξ^0	2570	½	neutre	$\overline{\Xi^0}$		$2{,}9 \times 10^{-10}$	$\Lambda^0 + \pi^0$

Notes :

– il existe deux K^0 qui se distinguent pour leur temps de vie ;

– d'après *The World of Elementary Particles*, par Kenneth W. Ford, © Copyright 1963, Xerox Corporation et *Review of Particles Properties*, par Particle Data Group, Reviews of Modern Physics, **56** (1984).

Le physicien américain Murray Gell-Mann et le physicien israélien Yuval Ne'eman ont proposé indépendamment quelque chose de ce genre en 1961. Des groupes de particules étaient disposés selon des configurations bien symétriques qui dépendaient de leurs diverses propriétés, un arrangement que Gell-Mann a appelé la *voie octuple*, mais plus connu maintenant sous le sigle de SU(3). L'un de ces regroupements, en particulier, nécessitait une particule de plus pour son achèvement. Cette particule devait, pour s'intégrer au groupe, avoir une masse donnée ainsi que d'autres propriétés bien spécifiques. Tout ceci donnait une combinaison improbable, mais une particule (l'*oméga-moins*), qui avait précisément l'ensemble des propriétés prédites, a pourtant été détectée en 1964, et l'on en a détecté par la suite des douzaines. En 1971, on a détecté son antiparticule, l'*antioméga-moins*.

En répartissant les baryons en groupes et en établissant un tableau périodique subatomique, il y avait encore assez de particules pour inciter les physiciens à trouver quelque chose de plus simple et plus fondamental. Gell-Mann, qui s'était attelé à la tâche de trouver la façon la plus simple de rendre compte de tous les baryons par un nombre minimal de *particules sub-baryoniques* plus fondamentales, a proposé en 1964 la notion de *quark*. Il a choisi ce nom après avoir trouvé que trois quarks combinés suffisaient pour constituer un baryon et que les différentes combinaisons de ces trois quarks permettaient de retrouver tous les baryons connus. Il a emprunté ce terme au roman de James Joyce, *Finnegans's Wake*.

Pour rendre compte des propriétés connues des baryons, il fallait aux trois quarks des propriétés bien spécifiques. La propriété la plus étonnante était la charge électrique fractionnaire. Toutes les particules connues avaient soit aucune charge électrique, soit une charge électrique exactement égale à celle de l'électron (ou du positron), soit encore une charge électrique qui était exactement un multiple de celle de l'électron (ou du positron). En d'autres termes, les charges connues valaient 0, +1, –1, +2, –2, etc. Proposer des charges fractionnaires était si bizarre que l'idée de Gell-Mann s'est heurtée d'abord à une vive résistance. Ce n'est que parce que son idée parvenait à expliquer tant de choses que Gell-Mann a été finalement écouté, suivi, et récompensé par le prix Nobel de physique en 1969.

Gell-Mann commençait, par exemple, avec deux quarks, ceux que l'on appelle maintenant *quark up* et *quark down. Up* (« haut ») et *down* (« bas ») n'ont ici aucune signification littérale et ne sont rien de plus qu'une façon fantaisiste de désigner ces quarks. (Il faut se garder de considérer les scientifiques, surtout les jeunes, comme des machines intellectuelles privées d'âme et d'émotions. Ils sont généralement aussi facétieux et parfois aussi stupides que le romancier ou le garçon de bains moyens.) Il vaut mieux les appeler *quark u* et *quark d*.

Le quark u a une charge de + 2/3 et le quark d de –1/3. Il doit y avoir un *antiquark u* de charge –2/3 et un *antiquark d* de charge + 1/3.

Deux quarks u et un quark d auraient une charge +2/3 +2/3 –1/3, soit un total de + 1 et, combinés, constitueraient un proton. D'autre part, deux quarks d et un quark u, avec leurs charges –1/3 –1/3 +2/3 = 0, se combineraient en un neutron.

Un groupement de trois quarks donnerait donc toujours une charge totale entière. Ainsi, deux antiquarks u et un antiquark d auraient une charge totale –1 et formeraient un antiproton, tandis que deux antiquarks d

et un antiquark u, avec une charge totale nulle, formeraient un antineutron.

De plus, les quarks seraient si fortement liés par l'interaction nucléaire que les chercheurs ont été jusqu'à présent totalement incapables de briser les protons ou les neutrons en quarks séparés. En fait, on pense que l'attraction entre les quarks augmente avec la distance, de sorte qu'il n'y a aucun moyen de décomposer un proton ou un neutron en quarks constituants. S'il en est ainsi, les charges fractionnaires peuvent bien exister, mais elles ne pourront jamais être détectées, ce qui rend la suggestion iconoclaste de Gell-Mann un peu plus facile à assimiler.

Mais ces deux quarks ne suffisent pas à rendre compte de tous les baryons et tous les mésons (qui sont des combinaisons de *deux* quarks). Gell-Mann avait dès l'origine proposé un troisième quark, maintenant appelé le *quark s*. On dit parfois que le *s* vient de « sideways » (« de côté », pour changer de haut et bas) ou, plus souvent, de « strangeness » (« étrangeté »), car ce quark était nécessaire pour rendre compte de la structure de certaines particules dites *étranges*, au sens où leur durée de vie était beaucoup plus longue que ce que l'on attendait.

En étudiant l'hypothèse des quarks, les physiciens ont finalement décidé qu'ils devaient exister par paires. S'il y avait un quark s, il lui fallait un compagnon qu'ils ont appelé *quark c*. (Le *c* n'est pas là pour « compagnon », mais pour « charme ».) En 1974, deux physiciens américains, Burton Richter et Samuel Chao Chung Ting, travaillant indépendamment avec de très hautes énergies, ont détecté des particules dont les propriétés nécessitaient le quark c. (On dit de ces particules qu'elles ont du charme.) Cette découverte leur a valu les honneurs et avantages du prix Nobel de physique 1976.

Les paires de quarks ont des saveurs qui, d'une certaine façon, correspondent aux saveurs des leptons. Chaque saveur de quark comporte quatre membres, par exemple le quark u, le quark d, l'antiquark u et l'antiquark d, tout comme chaque saveur de lepton a quatre membres, par exemple l'électron, le neutrino, l'antiélectron et l'antineutrino. Il y a dans les deux cas trois saveurs connues : l'électron, le muon et le tau pour ce qui est des leptons, et pour les quarks, u-d, s-c et finalement t-b. Le *quark t* et le *quark b* sont les formes abrégées de *top* (« dessus ») et *bottom* (« fond ») dans la terminologie usuelle, ou de *truth* (« vérité ») et *beauty* (« beauté ») chez les fantaisistes. Les quarks, comme les leptons, semblent être des particules ponctuelles et donc fondamentales et sans structure (mais comment en être sûr après avoir été abusé successivement par l'atome et par le proton ?). Et dans les deux cas il peut s'avérer y avoir un nombre infini de saveurs si l'on dispose de plus en plus d'énergie à dépenser pour les détecter.

Il y a une énorme différence entre les leptons et les quarks : les leptons ont des charges entières ou nulles et ne se combinent pas, tandis que les quarks ont des charges fractionnaires et n'ont apparemment d'existence qu'en combinaison.

Les quarks se combinent en obéissant à certaines règles. Chaque saveur de quark est disponible dans trois états d'une propriété que ne possèdent pas les leptons. Cette propriété s'appelle (métaphoriquement seulement) la *couleur*, et les trois états sont dits *rouge, bleu* et *vert*.

Lorsque les quarks se mettent à trois ensemble pour faire un baryon, il faut un quark rouge, un bleu et un vert, pour que la combinaison soit

incolore, ou *blanche*. (C'est la raison du rouge, du bleu et du vert pour désigner les états de quarks, car dans notre monde à nous, comme sur les écrans de télévision, cette combinaison donne du blanc.) Lorsque les quarks se mettent à deux pour fabriquer un méson, l'un est d'une couleur donnée et l'autre de l'anticouleur correspondante, de sorte que le mélange est encore blanc. (Les leptons n'ont pas de couleur, ils sont blancs au départ.)

L'étude des combinaisons de quarks telles que la couleur du produit final soit, comme les charges fractionnaires, indétectable, est appelée la *chromodynamique quantique, chromo* venant du mot grec qui signifie « couleur ». (Ce titre rappelle une théorie moderne des interactions électromagnétiques, couronnée de succès, appelée *électrodynamique quantique*.)

Lorsque les quarks se combinent, ils tiennent ensemble au moyen d'une particule qui passe continuellement de l'un à l'autre. Pour des raisons évidentes on appelle cette particule d'échange le *gluon*. Les gluons ont aussi une couleur, ce qui complique le tableau, et ils peuvent même se coller ensemble pour former une *boule de glu*.

Même si l'on ne peut démonter les hadrons pour obtenir des quarks isolés (deux dans le cas des mésons, trois pour les baryons), on dispose quand même de preuves indirectes de l'existence des quarks. On pourrait peut-être fabriquer des quarks à partir du vide en concentrant suffisamment d'énergie dans un petit volume, par exemple en précipitant face à face des faisceaux d'électrons et de positrons de très haute énergie (comme cela a permis de fabriquer des taus).

Les quarks produits de cette façon se combineraient instantanément en hadrons et antihadrons qui jailliraient dans des directions opposées. Avec *assez* d'énergie, il y aurait trois jets formant un trèfle à trois feuilles, l'hadron, l'antihadron et un gluon. On a déjà fait des trèfles à deux feuilles, et c'est en 1979 que les expériences ont donné un début de trèfle à trois feuilles rudimentaire. Ces observations constituent une importante confirmation de la théorie des quarks.

Les champs

Toute particule douée de masse est la source d'un champ gravitationnel qui s'étend jusqu'à l'infini dans toutes les directions, l'intensité du champ décroissant comme l'inverse du carré de la distance à la source.

L'intensité du champ créé par des particules individuelles est très petite, si petite qu'à toute fin pratique on peut négliger ce champ, lorsqu'on étudie les interactions entre particules. Néanmoins, il n'y a qu'une seule sorte de masse, et il semble que l'interaction gravitationnelle entre deux particules soit toujours attractive.

De plus, lorsqu'un système est constitué d'un grand nombre de particules, le champ gravitationnel semble être la somme des champs individuels de toutes les particules. Un corps comme le Soleil ou la Terre a, extérieurement, le champ que l'on attendrait d'une particule ayant la masse du corps et concentrée en son centre d'inertie. (Ceci n'est strictement exact que pour un corps parfaitement sphérique et de densité uniforme,

ou tout au moins distribuée selon une symétrie sphérique ; c'est pratiquement le cas pour des objets comme le Soleil ou la Terre.)

Il en résulte que le Soleil, et dans une moindre mesure la Terre, ont des champs gravitationnels d'une intensité énorme, qu'ils peuvent interagir, s'attirer mutuellement et rester fermement unis, même s'ils sont séparés par une distance de 150 millions de kilomètres. Des systèmes galactiques peuvent tenir ensemble bien qu'étalés sur des distances de millions d'années-lumière, et si jamais l'Univers se contracte à nouveau, ce sera à cause de l'attraction de la gravité sur une distance de l'ordre du milliard d'années-lumière.

Toute particule douée d'une charge électrique est la source d'un champ électromagnétique qui s'étend jusqu'à l'infini dans toutes les directions et dont l'intensité décroît comme l'inverse du carré de la distance à la source. Toute particule possédant à la fois une masse et une charge électrique (et il n'existe pas de charge électrique sans masse) est source de ces deux champs.

L'INTERACTION ÉLECTROMAGNÉTIQUE

Pour une particule donnée, le champ électromagnétique est des myriades de fois plus intense que le champ gravitationnel. Mais il existe deux sortes de charge électrique, positive et négative, et le champ électrique peut être aussi bien répulsif qu'attractif. Lorsque les deux sortes de charge sont en quantités égales, elles tendent à se neutraliser mutuellement, et le champ électromagnétique extérieur décroît plus vite que l'inverse du carré de la distance. C'est le cas des atomes ordinaires qui, étant constitués de nombres égaux de charges positives et négatives, sont électriquement neutres.

En présence d'un excès de l'une ou l'autre des deux sortes de charge, on a évidemment un champ électromagnétique, mais du fait de l'attraction mutuelle des charges opposées, celles-ci tendent à se neutraliser et il ne reste que le champ créé par l'excès de charge ; c'est la raison pour laquelle les champs électromagnétiques ont une intensité qui ne peut se comparer à celle des champs gravitationnels, dès que l'on a affaire à des corps plus massifs que les petits astéroïdes. C'est pourquoi Isaac Newton, en ne considérant que les interactions gravitationnelles *seules*, put élaborer une explication des mouvements des corps dans le système solaire, valable même pour les étoiles et les galaxies.

On ne peut ignorer complètement les interactions électromagnétiques car elles jouent un rôle dans la formation du système solaire, dans les transferts de moment angulaire du Soleil aux planètes, et probablement dans quelques phénomènes étranges parmi les petites particules qui constituent les anneaux de Saturne, mais il s'agit là de petits raffinements.

Tout hadron (un méson, ou un baryon, ou leurs quarks constitutifs) est la source d'un champ qui s'étend dans toutes les directions, mais dont l'intensité décroît si rapidement avec la distance qu'il ne se fait guère sentir au-delà d'une distance de l'ordre du diamètre du noyau atomique. Bien qu'il soit important dans un noyau ou lorsque deux particules se frôlent rapidement à une distance de l'ordre des dimensions nucléaires, on peut purement et simplement ignorer ce champ à plus grande distance. Ce champ ne joue aucun rôle dans les mouvements des corps célestes, mais devient par contre important pour les événements qui se déroulent au cœur des étoiles par exemple.

Les leptons aussi sont la source d'un champ qui ne se fait sentir que sur des distances à l'échelle nucléaire. La portée de ce champ est encore plus faible que celle du champ hadronique. Ce sont tous deux des champs nucléaires bien que très différents, non seulement par le type de particules auxquelles ils sont associés, mais aussi par leur intensité. Le champ hadronique, de particule à particule, est 137 fois plus intense que le champ électromagnétique. Le champ leptonique est environ 10^{11} fois plus faible que le champ électromagnétique. Le champ hadronique est donc habituellement appelé *interaction forte*, et le champ leptonique *interaction faible*. (N'oublions pas que l'interaction faible, bien que faible par rapport aux interactions forte et électromagnétique, est encore environ 10^{28} fois plus forte que l'interaction gravitationnelle.)

Ces quatre interactions rendent compte, pour autant que nous le sachions, de tous les comportements des particules et donc de toute mesure que l'on peut effectuer. Il n'y a aucun indice d'existence d'une éventuelle cinquième interaction. (Bien sûr, dire que ces interactions rendent compte de tout ce qui est mesurable est très loin de signifier que l'on comprenne à l'heure actuelle tous les comportements mesurables. Le fait de savoir qu'une équation mathématique complexe a une solution ne signifie pas nécessairement que l'on puisse trouver cette solution.)

Le premier traitement mathématique de l'interaction faible est dû à Fermi et remonte à 1934, mais pendant des dizaines d'années celle-ci est restée la moins bien connue des quatre interactions. Ces quatre interactions devraient par exemple avoir des particules d'échange correspondantes. Il y a le photon pour l'interaction électromagnétique, le graviton pour l'interaction gravitationnelle, le pion pour l'interaction forte au niveau du neutron et du proton, et le gluon pour l'interaction forte au niveau des quarks. Il devrait aussi exister une particule de ce genre, appelée la *particule W*, pour l'interaction faible (« Weak » en anglais) ; mais cette particule est restée introuvable pendant près d'un demi-siècle.*

LES LOIS DE CONSERVATION

Il nous faut aussi parler des lois de conservation qui fixent les règles permettant de savoir quelles sont les interactions possibles entre particules données et quelles sont les interactions qui ne jouent aucun rôle, et plus généralement de juger de ce qui peut se produire dans l'Univers et de ce qui ne peut pas arriver. Sans les lois de conservation, les événements de l'Univers seraient anarchiques et totalement incompréhensibles.

Les physiciens nucléaires disposent d'environ une douzaine de lois de conservation. Parmi elles on trouve les lois de conservation habituelles de la physique du XIXe siècle : la conservation de l'énergie, la conservation de l'impulsion, la conservation du moment cinétique, et la conservation de la charge électrique. Et puis il y a des lois de conservation moins familières : la conservation de l'étrangeté, la conservation du nombre baryonique, la conservation du spin isotopique, etc.

---

* Signalons qu'il en est encore ainsi, et pour longtemps, en ce qui concerne le graviton. (*N.D.T.*)

Il semble que les interactions fortes satisfont toutes ces lois de conservation, et au début des années cinquante les physiciens admettaient le caractère universel et irrévocable de ces lois. Mais ce n'était pas le cas. Certaines lois de conservation sont enfreintes par les interactions faibles.

C'est la loi de *conservation de la parité* qui fut d'abord mise en pièces. La parité est une propriété purement mathématique difficile à décrire en termes concrets ; contentons-nous de savoir que cette propriété concerne une fonction mathématique associée au comportement ondulatoire d'une particule et à sa position dans l'espace. La parité a deux valeurs possibles, *paire* et *impaire*. Ce qui compte, c'est que la parité était considérée comme une propriété fondamentale qui, comme l'énergie ou l'impulsion, était soumise à une loi de conservation : dans toute réaction ou transformation, la parité doit être conservée. Cela signifie que lorsque des particules interagissent pour en former de nouvelles, les parités des deux membres de l'équation doivent s'équilibrer comme les nombres de masse, les numéros atomiques, ou les moments cinétiques, croyait-on.

Je vais illustrer la chose. Si une particule de parité impaire interagit avec une particule dont la parité est paire pour former deux autres particules, l'une des nouvelles particules doit être impaire et l'autre paire. Si deux particules impaires donnent deux autres particules, celles-ci doivent être toutes les deux paires ou toutes les deux impaires. Si une particule paire se désintègre en deux particules, elles sont toutes les deux paires ou toutes les deux impaires. S'il se forme trois particules, elles sont toutes les trois paires, ou alors l'une est paire et les deux autres impaires. (Ceci peut être plus facile à voir si vous considérez les nombres pairs et impairs qui suivent des règles similaires. Un nombre pair, par exemple, ne peut être que la somme de deux nombres pairs ou de deux nombres impairs, mais jamais la somme d'un nombre impair et d'un nombre pair.)

Les ennuis avaient commencé lorsque l'on s'était aperçu que les mésons K se désintégraient parfois en deux mésons pi (ce qui, puisque le méson pi a une parité impaire, faisait au total une parité paire) et d'autres fois en trois mésons pi (soit une parité totale impaire). Les physiciens avaient conclu qu'il existait deux types de mésons K, l'un de parité paire et l'autre de parité impaire, qu'ils avaient nommés respectivement *théta* et *tau*.

Mais ces deux mésons étaient, à l'exception de leur parité, identiques en tout point ; ils avaient même masse, même charge, même stabilité, et quantité d'autres propriétés semblables. Qu'il puisse exister deux particules avec exactement les mêmes propriétés paraissait difficilement croyable. Se pouvait-il que ces deux particules n'en fussent en fait qu'une seule et était-ce dans l'idée de conservation de la parité que quelque chose n'allait pas ? C'est précisément ce que proposèrent en 1956 deux jeunes physiciens chinois qui travaillaient aux États-Unis, Tsung Dao Lee et Chen Ning Yang. Selon eux, bien que la conservation de la parité fût assurée dans les interactions fortes, elle pouvait être mise en défaut dans les interactions faibles, comme celles qui intervenaient dans la désintégration des mésons K.

En développant mathématiquement cette hypothèse, il leur était apparu que si la conservation de la parité s'effondrait, alors les particules impliquées dans les interactions faibles devaient avoir un *sens*, idée qui avait été avancée pour la première fois par le physicien hongrois Eugene Wigner. Je m'explique.

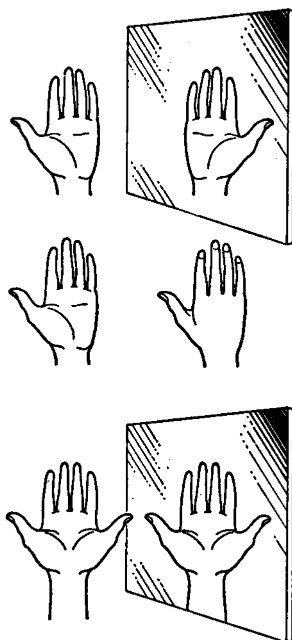


Figure 7.9. Illustration, avec des mains naturelles et imaginaires, de la réflexion dans un miroir d'objets asymétriques ou symétriques.

Votre main droite et votre main gauche sont opposées. L'une peut être considérée comme l'image de l'autre dans un miroir : la main droite vue dans un miroir a l'aspect de la main gauche. Si les mains étaient symétriques en tout point, leurs images dans un miroir ne seraient pas différentes des originaux, et il n'y aurait dans le principe aucune distinction du type main « gauche » et main « droite » (figure 7.9). Cela dit, appliquons ce principe à un ensemble de particules qui émettent des électrons. S'il sort le même nombre d'électrons dans toutes les directions, les particules en question n'ont pas de « sens ». Mais la plupart des électrons ont une direction d'émission préférentielle – disons vers le haut, plutôt que vers le bas –, alors les particules ne sont pas symétriques. Ces particules ont un « sens » ; si nous observons cette expérience dans un miroir, la direction préférentielle est inversée.*

Il faut donc observer un ensemble de particules qui émettent des électrons par interaction faible (disons des particules qui se désintègrent par émission bêta) et voir si les électrons sortent dans une direction préférentielle. Lee et Yang proposèrent cette expérience à une physicienne de l'université Columbia, Chien-Shiung Wu. Wu réalisa les conditions nécessaires. Il fallait aligner tous les noyaux émetteurs d'électrons dans la même direction si l'on voulait pouvoir détecter une direction d'émission

* Cette explication est un peu rapide. La lectrice intéressée consultera agréablement l'ouvrage de Martin Gardner, *L'Univers ambidextre* (Le Seuil, « Science ouverte », 1985). (*N.D.T.*)

privilégiée ; ceci était assuré par un champ magnétique, et la source était maintenue à une température proche du zéro absolu.

L'expérience donna la réponse en l'espace de quarante-huit heures. Les électrons étaient émis asymétriquement. La conservation de la parité était violée dans les interactions faibles. Le méson thêta et le méson tau n'étaient qu'une seule et même particule qui se désintégrait avec une parité paire dans certains cas et une parité impaire dans les autres cas. D'autres expérimentateurs confirmèrent bientôt la chute de la parité et les physiciens théoriciens Lee et Yang reçurent le prix Nobel de physique en 1957 pour leur audacieuse conjecture.

Si la symétrie est brisée par les interactions faibles, elle est peut-être brisée dans d'autres circonstances. Après tout, l'Univers dans sa totalité est peut-être gaucher (ou droitier). Ou alors il pourrait y avoir deux univers, un gaucher et un droitier, l'un fait de matière et l'autre d'antimatière.

Les physiciens considèrent maintenant les lois de conservation en général avec un nouveau cynisme. N'importe laquelle d'entre elles pourrait être valide sous certaines conditions et pas dans d'autres.

La parité, après sa chute, a été combinée avec la *conjugaison de charge*, une autre propriété mathématique attribuée aux particules subatomiques, qui contrôle leur statut de particule ou d'antiparticule ; les deux ensemble sont qualifiées de *conservation de PC*, une conservation plus fondamentale et plus générale que les conservations de la parité (P) ou de la conjugaison de charges (C) isolées. (Cette situation n'est pas sans précédent. Comme nous le verrons dans le chapitre suivant, la loi de conservation de la masse a, en son temps, cédé la place à la loi plus fondamentale de la *conservation de la masse-énergie*.)

Mais la conservation de PC s'est révélée aussi inadéquate. En 1964, deux physiciens américains, Val Logsden Fitch et James Watson Cronin ont montré que la conservation de PC était aussi, en de rares occasions, violée par les interactions faibles. La question du sens du temps (T) est alors venue se greffer, et on parle maintenant de la *symétrie PCT*. Fitch et Cronin ont partagé le prix Nobel de physique pour leurs travaux en 1980.

UNE THÉORIE DES CHAMPS UNIFIÉS

Pourquoi devrait-il y avoir quatre champs différents, quatre façons différentes d'interagir pour les particules ? Il pourrait bien sûr y en avoir un nombre quelconque, mais le désir de simplicité est bien ancré dans l'attitude des scientifiques. S'il y en a quatre (ou quelque autre nombre), ne pourrait-il pas s'agir d'aspects différents d'un champ unique, d'une seule interaction ? Si tel est le cas, la meilleure façon de s'en assurer serait de trouver une relation mathématique qui les exprimerait toutes et qui éclairerait un aspect de leurs propriétés resté dans l'ombre. A titre d'exemple, Maxwell trouva il y a une centaine d'années un ensemble d'équations mathématiques qui décrivaient aussi bien l'électricité que le magnétisme, et il montra qu'il ne fallait voir là que des aspects d'un seul phénomène que l'on appelle maintenant le *champ électromagnétique*. Ne pourrait-on aller un peu plus loin ?

Einstein avait commencé à chercher une *théorie des champs unifiés* à une époque où l'on ne connaissait que les champs électromagnétique et gravitationnel. Il y passa des dizaines d'années et échoua ; et pendant qu'il

cherchait, on découvrit les deux champs à courte portée, ce qui ne facilitait pas sa tâche.

Cependant, dans la fin des années soixante, le physicien américain Steven Weinberg et le physicien pakistanais Abdus Salam, travaillant indépendamment, ont inventé un traitement mathématique commun pour le champ électromagnétique et le champ faible, l'ensemble s'intitulant *champ électrofaible*. La méthode a été ensuite perfectionnée par le physicien américain Sheldon Lee Glashow, un ancien camarade de lycée de Weinberg. La théorie impliquait l'existence, pour l'interaction faible comme pour l'interaction électromagnétique, de termes de *courants neutres*, c'est-à-dire la possibilité d'interactions entre particules sans échange de charge électrique. Certaines de ces interactions étaient inconnues auparavant, et lorsqu'on se mit à les chercher, on découvrit qu'elles existaient, exactement comme prévu ; c'était une preuve convaincante apportée à la nouvelle théorie. Weinberg, Salam et Glashow ont partagé le prix Nobel de physique en 1979.

La théorie électrofaible donnait des détails sur les particules d'échange manquantes pour l'interaction faible (il y avait un demi-siècle que l'on cherchait en vain ces particules). Il ne devait pas y avoir juste une particule W, mais trois particules, une W^+, une W^- et quelque chose appelé Z^0 ou, en d'autres termes, une particule positive, une négative et une neutre. On pouvait en outre spécifier certaines de leurs propriétés, si la théorie électrofaible était correcte. Par exemple, ces particules devaient avoir environ 80 fois la masse du proton, ce qui expliquait la difficulté de les mettre en évidence. Il fallait des énergies énormes pour les créer et les rendre détectables. De plus, à cause de ces énormes masses, l'interaction faible avait une *très* courte portée et il était peu probable que deux particules s'approchent suffisamment pour que l'interaction fasse son office, ce qui expliquait l'extrême faiblesse de l'interaction faible comparée à la forte.

Mais en 1983, les physiciens ont eu à leur disposition l'énergie suffisante pour créer ces particules ; on les a détectées toutes les trois, et avec la masse prévue. La théorie électrofaible était bien établie.

Entre-temps, il était apparu à de nombreux physiciens que le même raisonnement mathématique qui semblait unifier les champs électromagnétique et faible pourrait aussi bien suffire, au prix de quelques complications, pour le champ fort. On en a proposé plusieurs versions. Si la théorie électrofaible est une théorie unifiée, alors une théorie qui inclut aussi le champ fort est une *théorie grande unifiée*, généralement abrégée en GUT (pour « Grand Unified Theory » ; il y a plusieurs théories candidates ; on parle de GUTs).

Pour amener le champ fort dans le giron d'une GUT, il faut supposer l'existence de particules d'échange ultra-lourdes, soit douze particules en plus des gluons. Etant plus massives que les W et les Z, elles sont encore plus difficilement détectables, et il n'y a actuellement aucun espoir de parvenir à en créer. Elles ont une portée encore plus courte que tout ce que l'on a envisagé jusqu'à maintenant. Le rayon d'action de ces particules d'échange ultra-lourdes pour le champ fort est inférieur à 10^{-15} fois le diamètre du noyau atomique.

A partir du moment où ces particules d'échange ultra-lourdes existent, il peut y en avoir qui passent d'un quark à l'autre dans un proton. Cet échange peut détruire l'un des quarks en le changeant en lepton. Avec

un quark en moins, le proton est devenu un méson destiné à se désintégrer et à finir en positron.

Mais, pour que l'échange ait lieu, il faut que les quarks (des particules ponctuelles) passent assez près l'un de l'autre pour être à portée de ces particules ultra-massives. Cette portée est si minuscule que, même dans le volume restreint d'un proton, une telle rencontre est peu probable. On a effectivement calculé que l'approche nécessaire se produit si rarement qu'un proton ne se désintégrerait qu'après une vie moyenne de 10^{31} années. Cette durée représente 10^{21} fois l'âge actuel de l'Univers. Il s'agit bien sûr d'une durée de vie *moyenne*. Certains protons pourraient vivre beaucoup plus longtemps et d'autres beaucoup moins. En fait, si l'on pouvait disposer de suffisamment de protons à la fois, il se produirait chaque seconde un certain nombre de désintégrations de protons. Par exemple, il y a peut-être trois milliards de désintégrations de protons par seconde dans l'ensemble des océans. (On a l'impression que c'est beaucoup, mais c'est en fait une quantité absolument négligeable, par rapport au nombre total de protons dans l'Océan.)

Les physiciens aimeraient bien détecter ces désintégrations et les distinguer parmi le grand nombre d'événements similaires qui peuvent se produire. La détection de ces désintégrations fournirait une preuve irréfutable en faveur des GUTs ; mais, comme dans le cas des ondes gravitationnelles, la sensibilité requise pour cette détection est à la limite du possible, et il s'écoulera sans doute longtemps avant que la question soit réglée d'une façon ou d'une autre.

On peut faire appel aux théories qui interviennent dans ces nouvelles unifications, pour calculer les détails du big bang qui a donné naissance à l'Univers. Il semble qu'au départ, il n'y avait qu'un seul champ et un seul type d'interaction, alors que l'Univers n'était encore âgé que de 10^{-42} seconde, qu'il était beaucoup plus petit qu'un proton et avait une température de 10^{36} degrés. Au cours de l'expansion de l'Univers, la température a baissé et les divers champs se sont « figés ».

Pour se représenter cela, on peut imaginer notre Terre si chaude qu'elle n'est rien qu'une boule gazeuse dans laquelle les différentes espèces d'atomes sont mélangées de façon homogène, de sorte que toutes les parties du gaz ont les mêmes propriétés. Mais au cours du refroidissement du gaz, diverses substances se sépareraient, d'abord sous forme liquide, puis solide, et on se retrouverait finalement avec une boule constituée de nombreuses substances séparées.

Pour l'instant, l'interaction gravitationnelle reste intraitable. Il semble qu'il n'y ait aucun moyen de l'intégrer dans le cadre du traitement mathématique mis au point par Weinberg et autres. L'unification qui défiait Einstein défie encore ses successeurs.

Malgré tout, les GUTs ont produit quelque chose de très intéressant. Les physiciens se sont demandé comment le big bang avait pu former un Univers granuleux, avec des galaxies et des étoiles. Pourquoi tout ne s'est-il pas simplement dispersé en une immense brume de gaz et de poussières dans toutes les directions ? Et puis, pourquoi l'Univers a-t-il juste la densité qui fait que l'on ne peut dire avec certitude s'il est ouvert ou fermé ? Il aurait pu être franchement ouvert (courbure négative) ou franchement fermé (courbure positive). Au lieu de cela il est pratiquement plat.

Un physicien américain, Alan Guth, a fait appel aux GUTs dans les années soixante-dix pour soutenir qu'après le big bang il y avait eu une

période initiale d'expansion, ou d'inflation, extrêmement rapide. Dans cet *Univers inflationniste,* la température a chuté si rapidement que les divers champs n'ont pas eu le temps de se séparer ni les particules de se former. Ce n'est que plus tard au cours du scénario que la différenciation a pu se produire, lorsque l'Univers était devenu très grand. D'où la platitude de l'Univers, et aussi sa granulosité. Le fait que les GUTs, des théories inventées seulement pour le besoin des particules, puissent expliquer deux énigmes qui concernent la naissance de l'Univers, témoigne fortement en faveur de leur exactitude.

L'Univers inflationniste ne résout quand même pas tous les problèmes, et plusieurs physiciens ont cherché à le ravauder de différentes manières pour améliorer l'accord entre les prédictions et la réalité ; il est encore bien tôt, mais beaucoup espèrent qu'une version des GUTs et l'inflation conviendront. Peut-être quelqu'un parviendra-t-il finalement à trouver une façon d'inclure l'interaction gravitationnelle dans la théorie, et l'unification sera enfin complète.

Chapitre 8

Les ondes

La lumière

Jusqu'à présent, je n'ai parlé que d'objets matériels – qu'il s'agisse d'objets extrêmement grands comme les galaxies, ou d'objets extrêmement petits comme les électrons, par exemple. Pourtant, il existe aussi des objets immatériels, des « immatériaux ». Le plus connu d'entre eux (qui est aussi celui dont les hommes ont le plus chanté les louanges), n'est autre que la lumière. Les premiers mots que prononce Dieu dans la Bible ne sont-ils pas : « Que la lumière soit ! » ? Et le Soleil et la Lune ne furent-ils pas créés afin « qu'ils soient des luminaires au firmament du ciel pour éclairer la Terre » ?

La véritable nature de la lumière est longtemps restée... obscure ; les savants de l'Antiquité et les érudits médiévaux se représentaient la lumière comme faite de particules émises par les objets incandescents, quand ce n'était pas par l'œil lui-même. Par contre, ils savaient déjà que la lumière se propage en ligne droite, qu'elle est réfléchie par les miroirs selon un angle égal à l'angle d'incidence, et qu'un rayon lumineux est réfracté (c'est-à-dire brisé) en passant de l'air dans l'eau, le verre, ou toute autre substance transparente.

LA VÉRITABLE NATURE DE LA LUMIÈRE

Quand la lumière pénètre dans le verre (ou toute autre substance transparente) de façon oblique (c'est-à-dire en faisant un angle avec la

verticale), elle est toujours réfractée de manière telle que son trajet se rapproche de la verticale. C'est le physicien hollandais Willebrord Snell qui, en 1621, a établi la relation exacte qui lie l'angle d'incidence à l'angle de réfraction. Snell n'a pas publié ses travaux et en 1637, le philosophe français René Descartes a redécouvert cette même loi.

Comme il a déjà été indiqué au chapitre 2, c'est Isaac Newton qui, en 1666, a effectué les premières véritables expériences portant sur la nature de la lumière. Dans une chambre obscure, Newton faisait arriver la lumière du Soleil filtrant à travers les fentes d'un volet sur l'une des faces d'un prisme de verre. Le faisceau, après avoir subi une première réfraction à l'entrée dans le prisme, émergeait en en subissant une deuxième. Comme les deux faces du prisme faisaient entre elles un certain angle, les rayons avaient subi deux déviations de même sens (contrairement à ce qui se passe dans le cas d'une lame de verre non prismatique, où les deux réfractions correspondent à des déviations de sens contraires – qui se compensent donc). Ayant recueilli le faisceau émergent doublement réfracté sur un écran, Newton constata qu'au lieu d'obtenir une simple tache blanche sur l'écran, il observait une large bande de couleurs, successivement : rouge, orange, jaune, vert, bleu, violet.

Newton en déduisit que la lumière blanche est un mélange de lumières différentes produisant sur l'œil des impressions colorées différentes. La bande de couleurs observée par Newton, pour réelle qu'elle lui parût, n'en était pas moins immatérielle, tout aussi immatérielle qu'un fantôme. Et de fait, Newton donna le nom de *spectrum* (qui signifie « spectre », « fantôme », en latin) à cette bande de couleurs.

Newton décréta alors que la lumière est faite de petites particules (des *corpuscules*), animées d'une vitesse extrêmement grande. Ainsi s'expliquait selon lui que la lumière aille en ligne droite et qu'elle dessine des ombres aux bords nets. De même, la réflexion sur un miroir pouvait s'expliquer par le rebondissement des corpuscules sur la surface du miroir ; quant à la brisure des rayons lumineux lors de la réfraction, Newton l'expliquait en supposant que les corpuscules ont, dans un milieu tel que le verre ou l'eau, une vitesse supérieure à celle qu'ils ont dans l'air.

Pourtant, certaines questions restaient sans réponse. Pourquoi les particules de lumière verte, par exemple, étaient-elles plus réfractées que celles de lumière jaune ? Comment expliquer que deux faisceaux de lumière puissent se croiser sans en être pour autant affectés, c'est-à-dire sans qu'il y ait collision entre les corpuscules ?

En 1678, le physicien hollandais Christiaan Huygens (également célèbre pour ses travaux en astronomie, et pour avoir construit la première horloge à balancier) émit une hypothèse allant à l'encontre de celle de Newton, et selon laquelle la lumière est constituée de petites ondes. Si l'on admet cette hypothèse, et si de plus, on admet que les ondes se propagent moins vite dans un milieu réfringent que dans l'air, alors il est possible d'expliquer pourquoi les lumières de différentes couleurs sont réfractées de façons différentes. Dans l'hypothèse ondulatoire, en effet, le taux de réfraction subie par la lumière est lié à la « longueur d'onde », longueur qui est en quelque sorte caractéristique de l'onde : plus cette longueur d'onde est courte, plus la réfraction est importante. D'où l'on doit conclure que la lumière violette a une longueur d'onde inférieure à celle de la lumière bleue, qui elle-même a une longueur d'onde plus courte que celle de la lumière verte, etc. Dans l'hypothèse de Huygens, cette différence de

longueur d'onde se traduit au niveau de l'œil par une différence de sensation colorée. De plus, si l'on admet l'hypothèse ondulatoire, il devient possible de comprendre pourquoi deux faisceaux peuvent se traverser sans en être affectés : les ondes sonores, ou bien les rides à la surface de l'eau, ne se croisent-elles pas sans perdre leur identité ?

Mais la théorie de Huygens n'était pas entièrement satisfaisante. Elle ne permettait pas d'expliquer pourquoi la lumière va en ligne droite et forme des ombres aux bords nets, ni pourquoi elle ne peut pas contourner les obstacles, tout comme le font les ondes sonores, ou les vagues à la surface de l'eau. Dans quel milieu se produisaient donc les vagues lumineuses ?

Pendant plus d'un siècle, les deux théories se sont affrontées. La *théorie corpusculaire* de Newton était, de loin, la plus en vogue, en partie parce qu'elle paraissait plus logique que sa rivale, mais en partie aussi parce qu'elle jouissait du crédit que lui valait le nom de Newton. C'est en 1801 que le médecin et physicien anglais Thomas Young fit basculer l'opinion dans l'autre camp en réalisant l'expérience suivante : on envoie un faisceau lumineux sur deux trous percés dans un écran opaque et on recueille la lumière qui en sort sur un écran placé au-delà. Si la lumière était réellement constituée de particules, on devrait observer sur l'écran une simple tache lumineuse plus brillante à l'endroit où les deux faisceaux se superposent. Or, ce n'est pas du tout ce que Young observait : l'écran faisait apparaître une série de franges alternativement brillantes et noires. Tout se passait comme si les deux faisceaux, en s'additionnant, produisaient de l'obscurité !

La théorie ondulatoire, par contre, permet d'expliquer facilement ce qui se passe : à l'endroit des franges brillantes, les ondes de l'un des faisceaux se trouvent renforcées par celles de l'autre ; autrement dit, les deux ondes sont *en phase* et leurs maxima se renforcent l'un l'autre. Les franges noires correspondent alors aux zones où les ondes sont comme des vagues, *en opposition de phase*, les creux de l'une venant compenser les crêtes de l'autre : au lieu de se renforcer, les deux ondes « interfèrent » et donnent finalement une intensité nulle.

On peut alors, connaissant la valeur de l'interfrange et la distance entre les deux trous, calculer la longueur d'onde des ondes lumineuses. Ces longueurs d'onde sont effectivement très courtes. La longueur d'onde de la lumière rouge, par exemple, est de l'ordre de 0,000075 centimètre. Aujourd'hui, on convient de mesurer les longueurs d'onde à l'aide d'une unité plus commode, l'angström, abrégé en Å ; un Å est égal à un millionième de centimètre. La longueur d'onde du rouge, à l'une des extrémités du spectre, vaut environ 7 500 Å, celle du violet, à l'autre extrémité, se situe aux environs de 3 900 Å, les autres couleurs du spectre correspondent à des longueurs d'onde situées entre ces deux valeurs limites.

Ces ordres de grandeur sont d'une importance extrême. En effet, si la lumière se propage en ligne droite et si elle découpe des ombres aux bords nets, c'est parce que les longueurs d'onde lumineuses sont beaucoup plus petites que les dimensions des objets ordinaires ; les ondes ne peuvent contourner un obstacle que si cet obstacle n'est pas trop grand comparé à leur longueur d'onde. Même les bactéries, par exemple, sont d'une taille nettement supérieure à celle des longueurs d'onde lumineuse ; c'est la raison pour laquelle la lumière permet d'en tracer le contour net lors de

leur observation au microscope. La lumière ne peut pas, par contre, contourner des objets qui, tels les virus, ont une taille du même ordre de grandeur que leur propre longueur d'onde.

Ce résultat est dû au physicien français Augustin Jean Fresnel ; en 1818, Fresnel a démontré que si l'on fait intérférer la lumière avec un objet suffisamment petit, elle le contourne effectivement. La lumière produit ce que l'on appelle une figure de *diffraction*. Un « réseau de diffraction » est obtenu en gravant des traits extrêmement fins sur une plaque. Ces traits se comportent comme autant d'obstacles minuscules dont les effets s'ajoutent. Comme la diffraction dépend de la longueur d'onde, on obtient un spectre coloré. Connaissant alors l'intensité diffractée correspondant à une couleur donnée (ou à une région du spectre donnée), ainsi que la distance qui sépare deux traits successifs du réseau, on peut, là encore, déterminer la longueur d'onde de la lumière.

Fraunhofer a été le premier à utiliser la méthode de la diffraction par un réseau pour déterminer les longueurs d'onde. (On a un peu tendance à oublier cette circonstance et à ne retenir que sa contribution à la découverte des raies spectrales.) Lorsque le physicien américain Henry Augustus Rowland eut inventé le réseau concave et mis au point des techniques permettant de fabriquer des réseaux à 10 000 traits par centimètre, le réseau remplaça définitivement le prisme dans les mesures spectroscopiques.

De tels exploits expérimentaux, joints au fait que Fresnel avait établi la théorie mathématique de la propagation de la lumière, laissaient penser que la théorie ondulatoire triomphait sur toute la ligne et que la théorie corpusculaire était à jamais disqualifiée. En effet, non seulement les ondes lumineuses existaient, mais on ne cessait d'accroître la précision sur la mesure de leurs longueurs d'onde. En 1827, le physicien français Jacques Babinet suggéra d'utiliser, à la place des étalons de longueur alors en vigueur, une longueur d'onde lumineuse – quantité inaltérable si jamais il en fût. Ce n'est cependant qu'aux environs de 1880 que cette idée put prendre forme grâce aux travaux du physicien américain d'origine allemande Albert Abraham Michelson, l'inventeur de l'*interféromètre*. Grâce à cet appareil, capable de mesurer des longueurs d'onde avec une précision jusqu'alors inconnue, Michelson établit la valeur de la longueur d'onde de la raie rouge du cadmium à 1/1 553 164 partie du mètre.

En fait, comme on le découvrit un peu plus tard, les éléments sont constitués de mélanges d'isotopes dont les raies spectrales ne sont pas rigoureusement identiques. Dès 1830, on était capable de mesurer la longueur d'onde de la raie orange de l'isotope 86 du krypton, avec une précision supérieure à la mesure de Michelson (le krypton est à l'état gazeux ; on peut donc diminuer l'élargissement de la raie en abaissant la température de manière à réduire l'agitation thermique des atomes).

En 1960, la Conférence internationale des poids et mesures décida d'adopter cette raie particulière de l'isotope 86 du krypton comme étalon de longueur. On gagnait ainsi trois ordres de grandeur sur la précision avec laquelle est défini l'étalon : alors que l'ancien mètre étalon ne pouvait être mesuré avec une précision supérieure au millionième, la lumière, elle, autorisait une précision d'un milliardième.

LA VITESSE DE LA LUMIÈRE

On ne peut guère en douter : la lumière se propage à une vitesse fantastique. Lorsqu'on éteint une lampe, par exemple, l'obscurité s'établit immédiatement, et partout, du moins à ce qu'il nous semble. Le son, lui, ne se propage pas aussi vite : lorsqu'on regarde quelqu'un couper du bois à une certaine distance, on perçoit très bien qu'il existe un intervalle de temps entre le moment où l'on voit la hache tomber sur le bois et celui où l'on entend le son de la cognée : manifestement, le son a mis un certain temps pour aller du bûcheron à notre oreille. De fait, la vitesse du son (dans l'air et au niveau de la mer) est de 332 m/s.

C'est Galilée qui, le premier, a essayé de mesurer la vitesse de propagation de la lumière. Galilée et son assistant, munis chacun d'une lanterne, se postaient au sommet de deux collines ; Galilée enlevait le cache qui recouvrait sa lanterne ; dès que l'assistant, sur l'autre colline, percevait la lumière, il enlevait lui aussi le cache qui recouvrait sa propre lanterne. Galilée a répété cette même expérience plusieurs fois, sur des distances de plus en plus grandes ; il avait posé a priori que le temps de réponse de son assistant était toujours le même, si bien que toute augmentation de l'intervalle séparant le moment où il découvrait sa lanterne et le moment où il percevait la lumière en provenance de l'autre colline devait correspondre à une augmentation du temps mis par la lumière pour parcourir la distance entre les deux collines de plus en plus éloignées.

L'hypothèse sur laquelle reposait cette expérience était parfaitement correcte ; malheureusement, la vitesse de la lumière est si élevée que Galilée ne put mettre en évidence aucune différence entre les divers intervalles de temps mesurés.

C'est en 1676 que l'astronome danois Olaus Roemer réalisa la première mesure de la vitesse de la lumière sur des distances astronomiques, précisément. Roemer, qui étudiait les éclipses des quatre satellites de Jupiter, remarqua que l'intervalle séparant deux éclipses successives augmentait au fur et à mesure que la Terre, dans son mouvement orbital, s'éloignait de Jupiter, et qu'il diminuait quand la Terre se rapprochait de Jupiter. Ces différences, visiblement liées aux variations de distance entre la Terre et Jupiter, devaient donc fournir un moyen de mesurer le temps mis par la lumière pour aller de Jupiter à la Terre. Partant d'une estimation grossière de la taille de l'orbite terrestre et de la valeur maximale de l'écart observé entre les intervalles séparant deux éclipses successives (écart qu'il convint d'identifier avec le temps mis par la lumière pour aller d'un bord à l'autre de l'orbite terrestre), Roemer calcula la valeur de la vitesse de la lumière. Le chiffre obtenu – 214 300 km/s – est remarquablement proche de la vraie valeur, surtout si l'on considère qu'il s'agissait d'un coup d'essai ; il était en même temps suffisamment grand pour provoquer l'incrédulité des contemporains.

Les résultats de Roemer ont été confirmés, un demi-siècle plus tard, par une mesure reposant sur un principe totalement différent. En 1728, l'astronome britannique James Bradley a montré que la position des étoiles varie au fur et à mesure que s'effectue le mouvement de la Terre sur son orbite et que cet effet, loin d'être un simple effet de parallaxe, doit en réalité être imputé au fait que la vitesse de la Terre sur son orbite représente une fraction, petite certes, mais non nulle, de la vitesse de la lumière. Pour faire comprendre cet effet, on utilise généralement l'image

d'un homme marchant sous la pluie et essayant de s'en préserver à l'aide d'un parapluie. Même si les gouttes de pluie tombent à la verticale, notre promeneur devra tenir son parapluie légèrement incliné vers l'avant ; et ce, parce qu'il marche à la rencontre des gouttes : plus il marche vite et plus il doit incliner son parapluie. On peut dire, de la même façon, que la Terre se déplace dans une pluie de lumière « qui tombe des étoiles » ; l'astronome doit modifier en conséquence l'inclinaison de son télescope au fur et à mesure que la Terre se déplace sur son orbite. Tel est ce qu'il est convenu d'appeler le phénomène d'*aberration des étoiles*. C'est l'observation de ce phénomène qui a permis à Bradley de mesurer la vitesse de la lumière et d'obtenir un résultat meilleur que celui de Roemer, mais encore inférieur de 5,5 % à la valeur vraie.

En définitive, c'est en raffinant l'idée initiale de Galilée qu'ont été obtenus les meilleurs résultats. En 1849, Armand Hippolyte Louis Fizeau construisit un dispositif ingénieux qui permettait d'envoyer la lumière d'un éclair lumineux se réfléchir sur un miroir distant de quinze kilomètres ; le temps mis par la lumière pour effectuer l'aller et retour n'excédait guère 1/20 000 de seconde ; mais Fizeau avait construit un autre dispositif, comportant une roue dentée placée sur le trajet de la lumière, qui lui permettait de mesurer des intervalles de temps aussi infimes. Pour certaines vitesses de rotation de la roue dentée, le rayon lumineux qui, à l'aller, passait entre deux dents successives, était arrêté au retour par la dent suivante et Fizeau, qui se trouvait derrière la roue, ne voyait rien ; il suffisait alors d'augmenter un peu la vitesse de rotation de la roue pour que la lumière au retour passe dans l'intervalle suivant, entre deux dents (voir figure 8.1). Si bien qu'en réglant et mesurant la vitesse de rotation

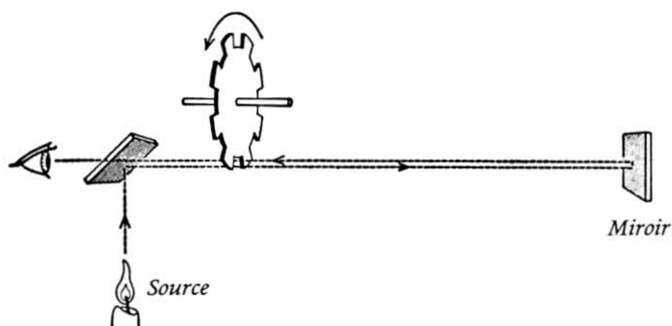


Figure 8.1. Dispositif utilisé par Fizeau pour mesurer la vitesse de la lumière. La lumière, d'abord réfléchie par le miroir semi-transparent placé près de la source, passe dans l'intervalle entre deux dents successives de la roue dentée, avant de se réfléchir sur le miroir (*à droite sur la figure*) et de revenir vers l'observateur. La roue dentée tourne à grande vitesse ; pour certaines vitesses de rotation, la lumière passe au retour comme à l'aller entre deux dents successives.

de la roue, Fizeau était capable de mesurer le temps mis par la lumière pour effectuer l'aller et retour, et donc d'en déduire une valeur de la vitesse de la lumière. Son résultat (315 300 km/s) était cette fois un peu trop grand par rapport à la valeur actuelle de 5,2 %.

L'année suivante, Jean Foucault (celui qui devait réaliser la fameuse expérience du pendule qui porte son nom, voir chapitre 4) reprit la même expérience en l'améliorant : il remplaça la roue dentée par un miroir tournant à vive allure et le temps mis par la lumière pour effectuer l'aller et retour était alors déduit de la mesure de la rotation du miroir (voir figure 8.2). La meilleure mesure de Fizeau, réalisée en 1862, est inférieure de 0,7 % à la valeur actuelle. Notons, en outre, que Fizeau utilisa le même dispositif pour mesurer la vitesse de propagation de la lumière dans divers liquides, et qu'il trouva des vitesses nettement inférieures à celles obtenue dans l'air, confirmant ainsi la théorie ondulatoire de Huygens.

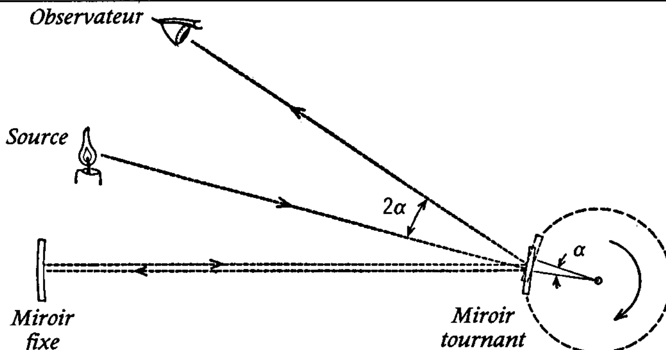
La précision sur la mesure de la vitesse de la lumière fut encore augmentée grâce aux travaux de Michelson qui, pendant près de quarante ans, reprit inlassablement l'expérience de Fizeau-Foucault, ne cessant de lui apporter des améliorations. Utilisant une pompe à vide mettant en œuvre des tuyaux de près de deux kilomètres de long, Michelson réalisa la première mesure de la vitesse de la lumière dans le vide. Son résultat, 299 796 km/s, ne diffère que de 0,006 %, par défaut, du résultat actuel. Il démontra également que la vitesse de la lumière dans le vide est la même quelle que soit la longueur d'onde.

En 1972, une équipe de chercheurs, sous la direction de Kenneth M. Evenson, établit le record de précision sur la mesure de la vitesse de la lumière dans le vide. La précision de cette mesure autorise l'emploi de la lumière comme moyen de mesure des distances (à vrai dire, on n'a pas attendu d'en arriver à une telle précision pour le faire).

LE RADAR

Considérons une impulsion lumineuse de courte durée se propageant en ligne droite et qui, après avoir heurté un obstacle, s'y réfléchit et revient à l'endroit d'où elle a été émise. Supposons de plus que la fréquence de cette impulsion soit à la fois suffisamment petite pour que l'onde puisse traverser le brouillard, la brume et les nuages, et suffisamment grande pour qu'elle se réfléchisse effectivement sur l'obstacle mentionné plus haut.

Figure 8.2. Méthode de Foucault. Ici c'est l'angle selon lequel le miroir a tourné durant le temps mis par la lumière pour effectuer l'aller et retour qui fournit une mesure de ce temps, et donc de la vitesse de la lumière.



Il se trouve que de telles conditions sont remplies pour un domaine de fréquences correspondant à des longueurs d'onde qui sont de l'ordre de quelques centimètres ou dizaines de centimètres (« radio-fréquences »). Connaissant alors le temps mis par l'onde pour effectuer l'aller et retour, on peut en déduire une estimation de la distance qui sépare le point d'émission de l'obstacle.

L'idée est simple et beaucoup de physiciens ont essayé, entre les deux guerres, de la mettre en œuvre ; c'est finalement le physicien écossais Robert Alexander Watson-Watt qui a mis au point le dispositif. Dès 1935 donc, on était en mesure de suivre un avion en recueillant son écho radio-fréquence. Le dispositif portait en anglais le nom de *radio detection and ranging* (le mot « ranging » indiquant ici la possibilité de déterminer la distance à laquelle se trouve l'avion), qui donna naissance au mot *radar* (*r.a.d.a.r*). (Il n'est pas rare de rencontrer dans le langage technique moderne des mots qui, tel radar, ont été forgés par contraction de plusieurs autres ; c'est là un procédé courant de fabrication de néologismes dans le domaine du vocabulaire technologique et scientifique.)

L'existence du radar ne fut révélée au monde que losqu'on apprit que c'était grâce au radar que les Anglais avaient été capables de repérer, lors de la bataille d'Angleterre, l'arrivée des avions allemands abordant les côtes du pays, malgré le brouillard et la nuit. C'est donc en partie grâce au radar que fut acquise la victoire sur les armées nazies.

Depuis la fin de la Seconde Guerre mondiale, les applications pacifiques du radar n'ont cessé de se multiplier. Ainsi, le radar permet de prévoir les orages en formation et il est donc utilisé en météorologie. C'est à cette occasion que l'on a mis en évidence des réflexions inexpliquées que pour cette raison on a baptisées du nom d'*anges*. En fait, il ne s'agit pas d'anges, mais simplement de bancs d'oiseaux migrateurs ; depuis lors, le radar sert aussi à étudier le vol migratoire des oiseaux.

Comme il a déjà été expliqué au chapitre 3, c'est également en étudiant les réflexions radar sur Vénus et Mercure que l'on a réussi à préciser le mouvement de rotation de ces planètes et, dans le cas de Vénus, à déterminer la nature de son sol.

DANS QUEL MILIEU LES ONDES LUMINEUSES SE PROPAGENT-ELLES ?

Bien que la nature ondulatoire de la lumière ne fût plus guère mise en doute, il restait néanmoins une question sans réponse. Comment la lumière peut-elle se propager dans le vide ? Par quel mécanisme se propage alors l'onde ?

Les autres types d'ondes, les ondes sonores par exemple, requièrent pour leur propagation un certain « milieu » ; l'impression de son que nous ressentons est alors due au mouvement de vibration des atomes et des molécules du milieu au travers duquel le son se propage. Il est impossible, par exemple, pour un observateur situé sur la Terre, d'entendre les explosions qui ont lieu sur la Lune, ou en n'importe quel autre endroit de l'espace intersidéral, et ce pour la simple raison que le son ne se propage pas dans le vide. Or, précisément, les ondes lumineuses, elles, semblent se propager plus facilement dans le vide que dans un milieu matériel : nous percevons la lumière qui vient des étoiles situées à des millions d'années-lumière, bien que rien ne vibre entre les deux !

Les physiciens de l'âge classique n'avaient jamais pu accepter complètement l'idée d'une « action à distance ». Newton, par exemple, était tracassé par l'idée que la force de gravitation puisse se faire sentir sans médiation matérielle, à travers l'espace. C'est pour tenter de trouver une explication à cette énigme qu'il en était venu à redonner vie à une idée autrefois émise par les savants grecs, celle d'un éther remplissant l'espace intersidéral ; Newton pensait que la force de gravitation devait probablement se propager grâce à l'éther. Le problème, dans l'esprit de Newton, ne se posait par contre pas dans le cas de la lumière, puisque la lumière, selon lui, était constituée de particules. Mais la question se posait de nouveau, maintenant que l'on était convaincu de la nature ondulatoire de la lumière.

Pour expliquer que la lumière puisse se propager à travers l'espace, les physiciens en vinrent alors à imaginer que la lumière, elle aussi, pourrait bien être véhiculée par l'intermédiaire de l'éther. Aussi commencèrent-ils à parler d'*éther luminifère* (c'est-à-dire porteur de lumière). Mais cette idée n'allait pas sans poser un certain nombre de problèmes. Les ondes lumineuses, en effet, sont des ondes *transverses*, c'est-à-dire qu'elles vibrent dans une direction perpendiculaire à leur direction de propagation – tout comme le font les ondes à la surface de l'eau, mais contrairement à ce qui se passe dans le cas du son, où les ondes correspondent à des vibrations longitudinales, le long de la direction de propagation. Or, on avait démontré que seul un milieu solide peut propager des ondes transverses (ainsi, les ondes à la surface de l'eau ne se propagent-elles précisément qu'à la surface ; elles ne peuvent pas pénétrer en profondeur). D'où il fallait conclure que l'éther était un milieu solide (et non pas liquide ou gazeux)... et présentant même une rigidité extrême : ce milieu, capable de propager des ondes à une vitesse aussi grande que celle de la lumière, devait être plus rigide que l'acier. Et pourtant ce milieu, pour rigide qu'il fût, devait quand même être capable de remplir tout l'espace et de pénétrer dans les interstices de la matière ordinaire, du moins de toute la matière transparente, liquide, gazeuse ou vitreuse, à travers laquelle la lumière se propage. Et pour couronner le tout, il fallait que ce milieu matériel solide, extrêmement rigide, ne donne lieu à aucune force de frottement et soit suffisamment élastique pour ne pas modifier le mouvement des plus petites planètes ni celui des battements de cils.

Néanmoins, en dépit de toutes ces difficultés théoriques, la notion d'éther était opératoire. Faraday (qui n'avait pas fait d'études théoriques poussées, mais qui était doué d'un sens physique hors du commun) avait élaboré le concept de *lignes de force* (lignes le long desquelles le champ magnétique a la même intensité) ; puis, en identifiant celles-ci aux distorsions élastiques de l'éther, il avait réussi à faire jouer à la notion d'éther un rôle fondamental dans l'explication des phénomènes.

Vers 1860, Clark Maxwell, admirateur passionné du travail de Faraday, entreprit de donner à l'idée de ligne de force conçue par Faraday un support mathématique. Il fut ainsi amené à introduire un ensemble de quatre équations qui, à elles quatre, permettaient de décrire la plupart des phénomènes électriques et magnétiques. Ces équations (publiées en 1864) non seulement mettaient en évidence le lien qui existe entre phénomènes électriques et phénomènes magnétiques, mais elles montraient également que les deux types de phénomènes sont indissolublement liés : tout champ électrique doit s'accompagner d'un champ magnétique qui lui est perpendiculaire (et réciproquement), si bien qu'en définitive

il n'existe qu'un seul type de champ : le champ électromagnétique. (Cette unification de deux champs en un seul a servi de modèle à tout le travail théorique du siècle suivant.)

En étudiant les conséquences prévues par ses équations, Maxwell s'aperçut qu'un champ électrique variable doit nécessairement produire un champ magnétique « induit », lequel à son tour « induit » l'existence d'un champ électrique variable, etc. Les deux champs jouent en quelque sorte à saute-mouton, et c'est ainsi que le champ se propage. D'où l'existence d'un rayonnement qui a toutes les caractéristiques d'une onde. En résumé, Maxwell avait prévu l'existence d'ondes électromagnétiques dont la fréquence devait être celle des fluctuations couplées du champ électrique et du champ magnétique.

Maxwell put même calculer la vitesse à laquelle une telle onde devait se propager. Pour cela , il lui avait fallu considérer le rapport entre la force s'exerçant entre deux charges électriques et celle s'exerçant entre deux pôles magnétiques. Or, la valeur ainsi trouvée était précisément celle de la vitesse de la lumière, ce qui ne pouvait pas être dû à un simple hasard. Maxwell suggéra donc que la lumière elle-même devait être considérée comme un *rayonnement électromagnétique*, un parmi beaucoup d'autres de longueurs d'onde plus petites ou plus grandes, tous ces rayonnements faisant intervenir l'éther comme milieu de propagation.

LES MONOPÔLES MAGNÉTIQUES

Les équations de Maxwell posent un problème qui, soit dit en passant, n'est toujours pas résolu. Elles font apparaître une parfaite symétrie entre les phénomènes magnétiques et les phénomènes électriques : ce qui vaut pour les uns doit valoir pour les autres. Pourtant, il semble bien qu'il existe entre les deux types de phénomènes une différence de nature fondamentale, et l'embarras dans lequel cette différence a plongé les physiciens n'a cessé de croître au fur et à mesure que l'on découvrait de nouvelles particules subatomiques. Toutes les particules qui portent une charge électrique portent soit des charges positives, soit des charges négatives, mais jamais les deux à la fois. Ne devrait-on pas s'attendre, par symétrie, à ne trouver que des particules qui se comportent soit comme un pôle + magnétique, soit comme un pôle – ? Pourtant, c'est en vain qu'on a cherché à déceler de tels *monopôles magnétiques*. Tout objet qui produit un champ magnétique, qu'il soit petit ou grand, qu'il s'agisse d'une galaxie ou d'une particule subatomique, se présente toujours comme une combinaison d'un pôle magnétique + et d'un pôle magnétique –.

En 1931, Dirac a montré, sur la base d'un raisonnement théorique, que s'il existe de tels monopôles magnétiques (plus même, s'il existe un seul monopôle magnétique, quelque part dans l'Univers), alors il faut obligatoirement que toutes les charges électriques soient des multiples entiers d'une même charge élementaire ; ce qui est effectivement le cas. Ne faut-il pas en conclure qu'il doit exister des monopôles magnétiques ?

En 1974, le physicien hollandais Gerard't Hooft et le Soviétique Alexandre Polyakov ont montré, chacun de leur côté, que dans le cadre des théories de grande unification, il doit effectivement exister des monopôles magnétiques et qu'ils doivent avoir une masse énorme. De tels monopôles devraient avoir une taille inférieure à celle du proton, mais

une masse dix à cent milliards de fois plus grande ; ces monopôles devraient avoir à peu près la même masse qu'une bactérie, mais être concentrés dans une région de l'espace qui ne dépasse pas la taille d'une particule subatomique.

De telles particules ne peuvent avoir été formées qu'au moment du big bang (car depuis lors, il n'y a jamais eu dans l'Univers de concentration d'énergie d'une telle ampleur). Ces particules gigantesques devraient avoir une vitesse d'environ 200 000 km/s. Le fait qu'elles possèdent une masse aussi énorme, joint à la faiblesse de leur taille, implique qu'elles puissent traverser la matière sans laisser trace de leur passage ; ce qui explique qu'elles soient si difficiles à mettre en évidence.

Le physicien Blas Cabrera de l'université de Stanford a monté une expérience destinée à mettre en évidence l'existence des monopôles : après avoir attendu quatre mois qu'un monopôle veuille bien se signaler dans son anneau supraconducteur, soigneusement isolé de toute perturbation magnétique, Cabrera a observé le 14 février 1982, à 13 h 53, une décharge électrique dont l'intensité correspondait pratiquement à celle que l'on s'attend à observer lors du passage d'un monopôle. On est en train, à l'heure actuelle, de monter d'autres expériences destinées à confirmer cette découverte ; pour le moment, cependant, on ne peut pas affirmer de façon certaine qu'un monopôle ait été effectivement détecté.

LE MOUVEMENT ABSOLU

Mais revenons à l'éther là où nous l'avons laissé, c'est-à-dire au sommet de la gloire. Pourtant, lui aussi devait bientôt connaître son Waterloo ; et ce, lors d'une expérience destinée à trancher au sein d'un débat tout aussi épineux que celui de l'action à distance, je veux parler du débat autour de la question du *mouvement absolu*.

Dans le courant du XIXe siècle, personne ne mettait plus en doute le mouvement des astres, qu'il s'agisse de la Terre, du Soleil, ou de tout objet placé dans l'Univers. Où, dans ces conditions, se trouvait le point de référence immobile, le centre absolument immobile par rapport auquel devait être défini le mouvement absolu, concept sur lequel reposait toute la théorie de Newton ? Restait pourtant une possibilité : Newton avait émis l'idée que la substance de l'espace (l'éther, selon toute probabilité) était immobile, ce qui permettait de parler d'un espace absolu. Si l'éther était immobile, il devait être possible de déterminer le mouvement *absolu* d'un objet en observant son mouvement *par rapport à l'éther*.

Aux environs de 1880, Michelson eut l'idée d'un montage ingénieux permettant de mettre en évidence précisément ce mouvement par rapport à l'éther. A supposer que la Terre soit en mouvement par rapport à l'éther, se disait Michelson, alors un rayon lumineux envoyé dans la direction de ce mouvement, puis réfléchi par un miroir, devrait avoir à parcourir une distance inférieure à celle parcourue par un rayon subissant le même genre de réflexion, mais dans une direction perpendiculaire à la précédente. Pour tester cette hypothèse, Michelson inventa un dispositif, appelé *interféromètre*. L'interféromètre de Michelson comporte un semi-miroir, c'est-à-dire un *miroir semi-réfléchissant*, qui laisse passer la moitié de l'intensité lumineuse et réfléchit l'autre moitié, ici à angle droit de la direction incidente. Les deux faisceaux issus du semi-miroir sont alors

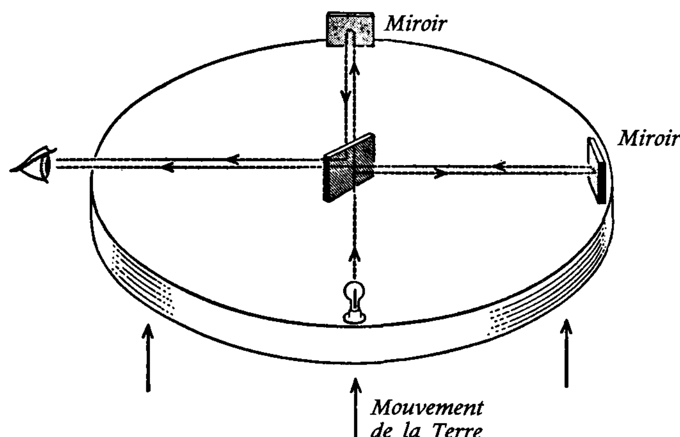


Figure 8.3. L'interféromètre de Michelson. Le miroir semi-transparent (*au centre*) partage le faisceau incident en deux parties : une partie réfléchie et une partie transmise. Si les deux miroirs totalement réfléchissants (*à droite en haut*) sont à des distances différentes du miroir semi-transparent, les faisceaux émergeant de l'appareil arrivent déphasés dans l'œil de l'observateur (*à gauche*).

réfléchis par deux vrais miroirs, puis de nouveau séparés en deux au niveau du semi-miroir (voir figure 8.3). Si les trajets suivis par les deux faisceaux ont des longueurs différentes, ils arrivent au niveau de l'œil en présentant une certaine différence de phase et l'on observe des interférences.

Cet appareil permet de mesurer des différences de longueur extrêmement faibles ; il est en fait tellement sensible qu'il permet d'observer la croissance d'une plante de seconde en seconde ou encore d'évaluer le diamètre d'une étoile qui, au télescope, apparaît comme un simple point.

L'idée de Michelson était la suivante : il se proposait de pointer son instrument dans des directions variables par rapport à la direction du mouvement de la Terre, et il espérait que l'effet de l'éther se traduirait par une variation de la différence de phase entre les deux « bras » de l'instrument.

En 1887, Michelson, aidé en cela par le chimiste américain Edward Williams Morley, réalisa une version particulièrement élaborée de son dispositif expérimental. L'appareil reposait sur une pierre, flottant elle-même sur une cuve pleine de mercure ; ainsi l'appareil pouvait-il être orienté facilement et sans secousse dans telle ou telle position à volonté. Michelson et Morley donnèrent à leur faisceau incident toutes les directions possibles par rapport à celle du mouvement de la Terre. En vain : aucune différence sensible ne se manifestait ! Les franges d'interférence étaient pratiquement inchangées, quelle que fût la direction dans laquelle l'appareil se trouvait orienté. (Précisons que des expériences récentes, plus sensibles encore que celle de Michelson et Morley, ont donné le même résultat négatif.)

La physique tremblait sur ses bases. De deux choses l'une : ou bien l'éther était entraîné par le mouvement de la Terre – ce qui n'avait guère de sens ; ou bien l'éther n'existait pas. Dans un cas comme dans l'autre, il n'existait ni espace absolu, ni mouvement absolu. La physique de Newton ne reposait plus sur rien ! La physique classique, celle de Newton, restait cependant encore valable dans le monde de tous les jours : les planètes effectuaient encore leurs révolutions selon les lois de la gravitation, et les objets terrestres obéissaient encore aux lois de l'inertie ou de l'égalité de l'action et de la réaction. Simplement, les explications classiques étaient incomplètes, et les physiciens devaient accepter l'idée que de nouveaux phénomènes allaient être découverts qui ne satisferaient pas aux lois de la physique classique. Les phénomènes observés étaient bien toujours les mêmes qu'avant, mais il allait maintenant falloir trouver d'autres théories, plus complexes et plus générales, pour les interpréter.

L'expérience de Michelson et Morley est certainement l'« expérience ratée » la plus importante de toute l'histoire de la physique. Michelson reçut le prix Nobel en 1907, pour l'ensemble de son œuvre. C'était le premier prix Nobel américain.

La théorie de la relativité

LES ÉQUATIONS DE LORENTZ-FITZGERALD

En 1983, le physicien irlandais George Francis FitzGerald produisit une théorie permettant de comprendre pourquoi l'expérience de Michelson et Morley avait donné un résultat négatif. Selon FitzGerald, la matière, lorsqu'elle est en mouvement, subit une contraction dans la direction de ce mouvement, et cette contraction est d'autant plus sensible que le mouvement est plus rapide. Le bras de l'interféromètre qui se trouve parallèle à la direction du mouvement de la Terre subit donc un raccourcissement, lequel compense exactement la différence de durée des trajets le long des deux bras. Qui plus est, dans la mesure où ce raccourcissement affecte tous les objets, y compris les organes des sens de l'observateur, il ne peut pas être mesuré par l'observateur qui est lui-même entraîné par la Terre en même temps que l'interféromètre. Tout se passe comme si la nature avait bâti une vaste conspiration destinée à nous empêcher de mettre en évidence le mouvement absolu : toute tentative destinée à mesurer les effets de ce mouvement est vouée à l'échec, dans la mesure où l'effet est immédiatement annulé par un effet inverse.

Tel est l'effet auquel on a donné le nom de *contraction de FitzGerald*, et pour lequel FitzGerald avait obtenu une formule quantitative. Un objet se déplaçant à environ 15 km/s (c'est à peu près la vitesse des fusées les plus rapides) devrait subir une contraction d'environ deux millièmes de millionième dans la direction de sa trajectoire. Mais à des vitesses réellement élevées cette contraction peut devenir tout à fait notable. Pour une vitesse de 150 000 km/s (la moitié de la vitesse de la lumière), elle atteint 15 % ; pour une vitesse de 280 000 km/s, elle avoisine 50 %. Autrement dit, une règle de vingt centimètres qui nous passerait sous le nez à une vitesse de 280 000 km/s nous semblerait n'avoir que dix

centimètres de long – à condition que nous ne soyons pas entraînés par le même mouvement que la règle et que nous ayons un moyen de mesurer la longueur d'une règle qui nous passerait devant les yeux à une telle vitesse ! La même règle se déplaçant à la vitesse de la lumière nous semblerait avoir une longueur nulle. Comme on ne peut imaginer qu'il existe des longueurs plus petites que zéro, on doit en déduire que la vitesse de la lumière dans le vide est la vitesse la plus grande à laquelle puisse se déplacer un objet dans l'Univers.

Le physicien hollandais Hendrik Antoon Lorentz poussa l'idée de FitzGerald un peu plus loin. Lorentz, à l'époque, travaillait sur les rayons cathodiques. Or, si l'on comprime la charge d'une particule chargée à l'intérieur d'un plus petit volume, la masse de la particule doit augmenter. Aussi, se dit Lorentz, une particule chargée qui, du fait de son mouvement, subirait la contraction de Lorentz, devrait aussi voir sa masse augmenter.

Lorentz calcula l'accroissement de masse ainsi produit et trouva un résultat du même type que celui obtenu par FitzGerald pour la contraction. A 150 000 km/s, un électron devrait subir une augmentation de masse de 15 %, une augmentation qui devrait passer à 100 % pour une vitesse de 280 000 km/s (la masse de l'électron doublerait) ; un électron qui aurait la vitesse de la lumière aurait une masse infinie. Là encore, il semblait bien que la vitesse de la lumière dans le vide représente une limite supérieure ; comment imaginer une masse supérieure à l'infini ? La contraction des longueurs de FitzGerald et l'augmentation de masse de Lorentz sont des effets liés, et les équations qui décrivent ces deux effets sont en général présentées ensemble ; elles portent le nom d'*équations de Lorentz-FitzGerald*.

Mais les variations de la masse d'un objet en fonction de sa vitesse peuvent être beaucoup plus facilement mesurées par un observateur immobile que les variations de sa longueur : le rapport entre la masse et la charge d'un électron, en effet, est directement donné par la mesure de la déflection subie par l'électron lorsqu'on lui fait traverser un champ magnétique. Si la masse de l'électron doit augmenter avec sa vitesse, par contre, il n'y a aucune raison de penser que sa charge doive être modifiée ; le rapport masse/charge doit donc augmenter avec la vitesse de l'électron, lequel doit ainsi avoir une trajectoire d'autant moins incurvée que sa vitesse est plus grande. Dès 1900, le physicien allemand Walter Kauffman était en mesure de montrer que le rapport masse/charge augmentait bien avec la vitesse, exactement selon la loi prévue par les équations de Lorentz-FitzGerald. Des mesures ultérieures plus précises vinrent confirmer ce point.

Il ne faut pas oublier, lorsqu'on parle de la vitesse de la lumière comme vitesse limite, qu'il s'agit de la vitesse de la lumière dans le vide (300 000 km/s). Dans un milieu matériel, la vitesse de la lumière est moins élevée : elle est égale à la vitesse de la lumière dans le vide divisée par la valeur de l'indice de réfraction du milieu considéré (l'*indice de réfraction* peut être considéré comme la mesure de la déviation subie par un rayon à sa traversée de la surface entre le vide et le milieu en question). Dans l'eau, qui a un indice de réfraction d'environ 1,3, la vitesse de la lumière est 1,3 fois plus petite que dans l'air (ou le vide) ; dans le verre, elle est réduite par un facteur 1,5 et dans le diamant par un facteur 2,4.

Rien ne s'oppose à ce qu'une particule subatomique ait dans un milieu transparent une vitesse supérieure à celle de la lumière dans ce même

milieu (mais pourtant inférieure à celle de la lumière dans le vide). Dans ce cas, la particule laisse derrière elle une traînée de lumière bleue, de la même façon qu'un avion voyageant à une vitesse supersonique émet une onde de choc.

Ce type de rayonnement a d'abord été observé par le physicien russe Pavel Alexeïevitch Tchérenkov (ou Čerenkov, selon les transcriptions), en 1934 ; l'explication théorique du phénomène a été donnée en 1937 par deux autres physiciens russes, Ilya Mikhailovitch Frank et Igor Ievghenievitch Tamm. Tchérenkov, Frank et Tamm ont reçu le prix Nobel en 1958 pour leurs travaux sur ce qu'il est convenu d'appeler *l'effet Tchérenkov*. Aujourd'hui on dispose de compteurs extrêmement sensibles à ce type de rayonnement, et c'est ainsi que l'on étudie certaines particules rapides telles celles qui constituent le rayonnement cosmique.

LE RAYONNEMENT DU CORPS NOIR ET LES QUANTA DE PLANCK

La physique n'avait pas fini de vaciller sur ses bases à la suite de l'expérience de Michelson et Morley que déjà se produisait une seconde explosion tout aussi violente. Cette fois-ci, c'est la question apparemment innocente du rayonnement produit par un corps chauffé qui allait mettre le feu aux poudres. (Bien que le rayonnement émis par un corps chauffé se présente essentiellement sous forme de lumière, les physiciens parlent à son propos de *rayonnement du corps noir* ; en fait, c'est parce qu'ils se réfèrent au rayonnement émis par un corps idéal qui absorberait complètement la lumière – sans rien réfléchir, ce qui est le propre des corps noirs – et qui, également, rayonnerait parfaitement, dans toutes les directions et quelle que soit la longueur d'onde considérée. Le physicien autrichien Josef Stefan avait déjà montré en 1879 que le rayonnement émis par un corps chauffé, si l'on tient compte de toutes les longueurs d'onde présentes dans ce rayonnement, ne dépend que de la température du corps, et en particulier pas de la nature de ce corps ; dans les conditions expérimentales idéales, ce rayonnement varie comme la puissance 4 de la température absolue du corps ; ce qui veut dire que si la température est multipliée par un facteur 2, la puissance rayonnée, elle, est multipliée par $2 \times 2 \times 2 \times 2$, soit 16 (*loi de Stefan*). On savait, par ailleurs, qu'au fur et à mesure que la température augmente, la longueur d'onde qui caractérise la lumière la plus fortement rayonnée décroît ; ce qui veut dire par exemple que si l'on chauffe un morceau de métal, il commence par rayonner principalement dans l'infrarouge (invisible), puis, au fur et à mesure que la température augmente, il rougeoie, d'abord faiblement puis fortement, avant de devenir successivement orange, puis jaune ; si le métal ne risquait pas de s'évaporer à de telles températures, on devrait finalement le voir tourner au bleu-violet.

En 1893, le physicien allemand Wilhelm Wien avait élaboré la théorie de ce phénomène et obtenu une expression mathématique pour la distribution spectrale du rayonnement du corps noir – c'est-à-dire la répartition de l'énergie rayonnée pour chaque domaine de longueur d'onde. La formule de Wien décrivait correctement ce que l'on observait du côté des courtes longueurs d'onde (dans la partie bleue du spectre), mais ne « collait » pas du tout avec les résultats expérimentaux à l'autre bout du spectre (dans le rouge). (Le prix Nobel fut décerné à Wien en 1911,

pour l'ensemble de ses travaux sur la théorie de la chaleur.) Par ailleurs, les physiciens anglais Lord Rayleigh et James Jeans avaient obtenu de leur côté et de façon indépendante une formule différente qui, si elle décrivait bien la distribution spectrale dans la partie rouge du spectre, ne « collait » pas du tout dans la partie violette. On était donc confronté à une situation où l'on disposait de deux théories qui chacune n'expliquait que la moitié du phénomène.

C'est à ce problème que le physicien allemand Max Karl Ernst Ludwig Planck entreprit de s'attaquer. Il s'aperçut alors que pour arriver à faire « coller » les formules avec les faits observés, il lui fallait introduire une notion radicalement nouvelle, à savoir que le rayonnement est constitué de « grains » d'énergie, un peu de la même manière que la matière est constituée de grains, les atomes. Planck baptisa ces unités élémentaires de rayonnement des *quanta* (pluriel du mot latin *quantum* qui signifie « combien »). Selon Planck, le rayonnement lumineux ne pouvait être absorbé que par nombres entiers de quanta. De plus, l'énergie de chaque quantum devait, selon lui, dépendre de la longueur d'onde du rayonnement considéré : plus la longueur d'onde est petite, plus le quantum correspondant est « énergétique » ; ou encore, l'énergie d'un quantum est inversement proportionnelle à la longueur d'onde du rayonnement considéré.

Il devenait alors possible d'établir un lien entre l'énergie d'un quantum et la fréquence du rayonnement. La fréquence est le nombre de vibrations émises en une seconde ; elle est inversement proportionnelle à la longueur d'onde (en effet, plus la longueur d'onde est petite, plus le nombre de vibrations émises en une seconde est grand). Puisque l'énergie d'un quantum et la fréquence du rayonnement sont toutes les deux inversement proportionnelles à la longueur d'onde, il s'ensuit que l'énergie et la fréquence sont proportionnelles entre elles. C'est ce que Planck traduisit en écrivant la fameuse relation :

$$e = h\,\nu$$

où e désigne l'énergie du quantum et ν (lettre grecque *nu*) la fréquence ; h est ce que l'on appelle la *constante de Planck*.

La constante de Planck a une valeur numérique extrêmement petite. Les quanta élémentaires de rayonnement sont de ce fait extrêmement petits et c'est pourquoi la lumière nous donne l'impression de quelque chose de continu – de la même façon que la matière ordinaire a l'air d'être continue. En somme, le rayonnement connaissait au début du XXe siècle le même sort que celui qu'avait connu la matière au début du XIXe siècle ; l'un et l'autre se révélaient être de nature discontinue.

La notion de quantum introduite par Planck permet d'établir la relation qui existe entre la température d'un corps et la distribution spectrale du rayonnement qu'il émet. En effet, un quantum de lumière violette qui a un contenu énergétique double de celui d'un quantum de lumière rouge ne peut être produit qu'en dépensant plus d'énergie (sous forme de chaleur). Les équations que l'on peut déduire en partant de l'hypothèse des quanta permettent d'expliquer la distribution spectrale du rayonnement du corps noir d'un bout à l'autre du spectre.

Mais finalement, la théorie de Planck devait apporter beaucoup plus que cela : l'explication du comportement des atomes, des électrons à

l'intérieur de l'atome, et des nucléons à l'intérieur du noyau de l'atome. Aujourd'hui on parle de *physique classique* pour désigner la physique d'avant Planck et de *physique moderne* pour désigner la physique postérieure à la découverte des quanta. Planck a reçu le prix Nobel en 1918.

EINSTEIN ET LA THÉORIE DES QUANTA

Lors de sa parution la théorie de Planck passa presque inaperçue ; son contenu était trop révolutionnaire pour que la majorité des physiciens l'acceptent d'emblée. Planck lui-même, d'ailleurs, semblait épouvanté de ce qu'il avait osé faire. Pourtant, cinq ans plus tard, un jeune physicien allemand d'origine suisse, nommé Albert Einstein, allait vérifier l'existence des quanta de Planck.

Le physicien allemand Philipp Lenard avait montré que, lorsque l'on envoie de la lumière sur certains métaux, la surface de ces métaux émet des électrons, comme si la lumière avait assez de force pour éjecter ces électrons en dehors des atomes du métal. Ce phénomène porte depuis le nom d'*effet photo-électrique*, et il valut le prix Nobel à son inventeur, Lenard, en 1905. Mais lorsque l'on commença à étudier le phénomène expérimentalement, on s'aperçut avec surprise que si l'on augmentait l'intensité de la lumière, les électrons éjectés n'en acquiéraient pas pour autant plus d'énergie. Par contre, si l'on changeait la longueur d'onde de la lumière incidente, alors les électrons éjectés en semblaient affectés : la lumière bleue, par exemple, donnait aux électrons émis une plus grande vitesse d'éjection que ne le faisait la lumière jaune. Une lumière bleue de faible intensité faisait sortir du métal moins d'électrons qu'une lumière jaune plus intense, mais ces électrons, peu nombreux, étaient animés d'une plus grande vitesse que les électrons plus nombreux émis sous éclairement en lumière jaune intense. Par ailleurs, dans le cas de certains métaux, la lumière rouge, quelle que soit son intensité, n'arrivait pas à extraire le moindre électron du métal.

Rien de tout cela ne s'expliquait si l'on s'en tenait aux termes de l'ancienne théorie de la lumière. Comment comprendre que la lumière bleue puisse faire quelque chose que la lumière rouge était incapable de faire ?

Einstein trouva la réponse à cette question dans la théorie de Planck même. Pour qu'un électron acquière une énergie qui lui permette de quitter le métal, il faut qu'il soit bombardé par un quantum suffisamment énergétique. Si l'électron n'est que faiblement lié au métal (c'est ce qui se passe dans le cas du césium), un quantum de lumière rouge, bien que faible, fera l'affaire. Mais si l'atome est plus fortement rattaché au métal, alors il faut utiliser de la lumière jaune, ou même de la lumière bleue, voire ultraviolette. Dans tous les cas, plus le quantum est énergétique, plus grande est la vitesse avec laquelle l'électron, une fois éjecté, quitte le métal.

La théorie des quanta expliquait donc de façon parfaitement limpide un phénomène que la conception préquantique de la lumière s'était montrée incapable d'interpréter. A partir de ce moment, les applications de la théorie quantique à d'autres phénomènes devinrent rapidement de plus en plus nombreuses. Einstein reçut le prix Nobel en 1921 pour sa théorie de l'effet photo-électrique.

Dans l'article (écrit en 1905) où il exposa sa théorie de la relativité restreinte – théorie qu'il avait élaborée à ses heures de liberté, une fois

son travail d'employé du bureau des brevets de Berne terminé –, Einstein fit état d'une nouvelle conception du monde fondée sur une généralisation de la théorie quantique. Il émit l'hypothèse que la lumière était constituée de quanta (le terme de *photon* qui, aujourd'hui, désigne les quanta lumineux ne fut introduit que plus tard par Compton, en 1928). Le concept de corpuscule lumineux se trouvait ainsi ressuscité. Mais en fait, il s'agissait d'un type de particule tout à fait nouveau ; car cette nouvelle particule présentait des propriétés à la fois ondulatoires et corpusculaires et manifestait les unes dans certains cas et les autres dans d'autres cas.

On a voulu voir là un paradoxe, et un certain discours mystique laisse entendre que la nature de la lumière est quelque chose qui dépasse l'entendement. Il n'en est rien, comme l'analogie suivante en convaincra le lecteur. Un homme peut très bien se manifester sous divers aspects de sa personnalité : soit comme mari, soit comme père, soit comme ami, soit comme homme d'affaires ; on ne s'attend pas à ce que le côté marital de sa personnalité se manifeste lorsqu'il traite des affaires, ni qu'il se comporte en homme d'affaires à l'égard de sa femme. Personne ne voit là le moindre paradoxe ; personne ne dira qu'il est « plusieurs hommes ».

De la même façon le rayonnement lumineux manifeste à la fois des propriétés corpusculaires et des propriétés ondulatoires. Dans certains cas, c'est le côté corpusculaire qui se manifeste le plus clairement ; dans d'autres cas, c'est l'inverse. Aux environs de 1930, Niels Bohr a donné une explication du fait que toute expérience destinée à mettre en évidence les propriétés ondulatoires du rayonnement ne permet pas d'en déceler la nature corpusculaire, et inversement ; selon Bohr, nous avons affaire soit à l'un soit à l'autre, soit à une onde soit à un corpuscule, mais jamais aux deux à la fois. C'est ce qu'il a appelé le *principe de complémentarité* : l'ensemble des propriétés, ondulatoires et corpusculaires, rend mieux compte de ce qu'est le rayonnement que les propriétés ondulatoires seules ou les propriétés corpusculaires seules.

La découverte de la nature ondulatoire de la lumière avait permis à l'optique et à la spectroscopie de se développer ; mais il avait fallu pour cela inventer, imaginer l'éther. La nouvelle conception d'Einstein n'abolissait pas les acquis du XIXe siècle (en particulier pas les équations de Maxwell) ; elle rendait simplement l'éther superflu : la lumière se propage dans le vide en vertu de ses propriétés corpusculaires, et l'idée d'éther que l'expérience de Michelson et Morley avait déjà sérieusement malmenée n'avait plus qu'à être oubliée.

Dans sa théorie de la relativité restreinte, Einstein avait introduit une autre idée nouvelle, à savoir que la vitesse de la lumière dans le vide est toujours la même, même si la source est en mouvement. Dans la conception classique de l'Univers, celle de Newton, un faisceau émis par une source en mouvement se dirigeant vers l'observateur doit sembler (à celui-ci) se propager plus vite qu'un faisceau émis par une source qui s'éloigne de l'observateur. Dans la conception einsteinienne, il n'en est rien : l'observateur voit les deux faisceaux se propager à la même vitesse. C'est en partant de cette hypothèse fondamentale qu'Einstein a retrouvé les équations de Lorentz-FitzGerald. De plus, Einstein a montré que l'augmentation de masse en fonction de la vitesse qui, chez Lorentz, ne concernait que les particules chargées, a une portée générale et vaut pour tous les objets quelle que soit leur nature. Enfin, Einstein a montré que pour un corps en mouvement, il y a non seulement contraction des

longueurs dans la direction du mouvement et augmentation de masse, mais également ralentissement du rythme d'écoulement du temps. Autrement dit, les règles se raccourcissent dans la mesure même où les horloges retardent.

LA THÉORIE DE LA RELATIVITÉ

La théorie de la relativité restreinte d'Einstein consiste essentiellement à affirmer qu'il n'existe ni espace absolu ni temps absolu. A première vue, cela semble absurde. Comment l'esprit humain peut-il s'y retrouver dans l'Univers s'il n'a rien de fixe à quoi se raccrocher ? Cela n'a pas d'importance, répond Einstein : tout ce dont nous avons besoin, c'est d'un système de référence par rapport auquel les événements qui se passent dans l'Univers puissent être repérés. N'importe quel *système de référence* (qu'il s'agisse de la Terre supposée immobile, ou du Soleil supposé immobile, ou de notre propre corps supposé également immobile) peut faire l'affaire ; aussi ce choix du système de référence peut-il s'effectuer selon des critères de commodité. Ainsi, par exemple, il est plus commode, pour calculer le mouvement des planètes, de choisir un système de référence dans lequel le Soleil est immobile plutôt qu'un système où c'est la Terre qui est immobile ; ceci dit, les deux choix sont tout aussi valables l'un que l'autre.

Ainsi donc, les mesures de temps et d'espace sont « relatives » à un certain système de référence, choisi de façon arbitraire. C'est pourquoi la théorie d'Einstein porte le nom de *théorie de la relativité*.

Un exemple fera mieux comprendre ce point. Supposons que nous observions, depuis la Terre, une planète inconnue (la planète X), semblable à la nôtre, de même masse et de mêmes dimensions, qui se déplace par rapport à la Terre à une vitesse de 280 000 km/s. A supposer que nous soyons capables de mesurer ses dimensions pendant le bref instant où elle passe devant nous, nous trouverions que sa dimension est réduite de moitié dans la direction de son mouvement. Elle aurait donc la forme d'un ellipsoïde (et non pas d'une sphère) et nous semblerait avoir une masse double de celle de la Terre.

Mais un observateur situé sur la planète X dirait que sa propre planète est immobile, que c'est la Terre qui passe devant lui à la vitesse de 280 000 km/s et qui semble avoir la forme d'un ellipsoïde et une masse double de celle de sa planète.

On est alors tenté de se demander qui a raison, quelle est la planète qui a *réellement* la forme d'un ellipsoïde et qui a *réellement* une masse double. Si cette question tracasse réellement le lecteur, qu'il se compare d'une part à une baleine et d'autre part à un coléoptère ; il est grand par rapport à l'un et petit par rapport à l'autre. Il est clair que la question « suis-je *réellement* petit ou grand ? » n'a pas de sens !

En dépit de toutes les conséquences plus ou moins déroutantes qu'elle implique, la théorie de la relativité explique les phénomènes observés dans l'Univers au moins aussi bien que les théories prérelativistes. De plus, elle explique certains phénomènes que la conception newtonienne était incapable d'expliquer. C'est pourquoi il faut considérer la théorie d'Einstein comme un raffinement de la théorie de Newton ; et c'est en ce sens que les physiciens ont opté pour Einstein plutôt que pour Newton. La

conception newtonienne du monde peut encore rendre des services : c'est une vision simplifiée, une approximation qui « marche » bien dans la vie courante, et même dans le domaine de l'astronomie ordinaire – pour placer des satellites en orbite par exemple. Mais il est des cas, lorsqu'on cherche à accélérer des particules dans un cyclotron, par exemple, où l'on ne peut pas ignorer la théorie d'Einstein, et en particulier l'augmentaion de masse subie par les particules à grande vitesse.

L'ESPACE-TEMPS ET LE PARADOXE DES HORLOGES

Le point de vue d'Einstein mélange les notions d'espace et de temps de façon telle que chacun de ces deux concepts, considéré isolément, perd toute signification. De fait, l'Univers est un espace à quatre dimensions, le temps devant être considéré comme l'une de ces quatre dimensions, une dimension qui n'a pas exactement le même statut que les trois autres, les dimensions d'espace ordinaires : longueur, largeur, hauteur. Cette union de l'espace et du temps en un seul espace à quatre dimensions est généralement appelée *espace-temps*. L'espace-temps est une notion introduite en 1907 par Hermann Minkowski, mathématicien allemand d'origine russe qui fut le professeur d'Einstein.

En dehors de ces aspects déroutants de l'espace et du temps, il est un autre trait de la théorie d'Einstein qui prête (encore maintenant) à controverse ; je veux parler du ralentissement des horloges. Selon Einstein, une horloge en mouvement bat moins rapidement qu'une horloge immobile. Plus généralement, tous les phénomènes qui présentent une évolution dans le temps évoluent moins vite lorsqu'ils sont entraînés et en mouvement que lorsqu'ils sont au repos. Autrement dit, le temps lui-même ralentit sa course. Aux vitesses ordinaires, cet effet est imperceptible ; mais une horloge animée d'une vitesse de 280 000 km/s semblerait, aux yeux d'un observateur immobile, mettre deux secondes pour battre la seconde. Et à la vitesse de la lumière, le temps cesserait de s'écouler...

Cet effet de dilatation du temps, comme on dit, est plus troublant que l'effet de contraction des longueurs ou d'augmentation de la masse. Si un objet est raccourci de moitié puis reprend sa taille ordinaire, ou si un objet double sa masse pour reprendre ensuite sa masse normale, il ne reste après coup aucune trace de ces modifications, et par conséquent ce phénomène ne peut donner lieu à controverse.

Le temps, par contre, est cumulatif. Imaginons qu'une horloge se trouve sur la planète X. Du fait de sa très grande vitesse, cette horloge nous semble battre à un rythme deux fois plus lent. Imaginons maintenant que l'horloge soit immobilisée (par rapport à nous) ; elle reprend son rythme ordinaire. Mais il reste une trace de ce qui s'est passé : l'horloge retarde d'une demi-heure ! Imaginons maintenant deux vaisseaux spatiaux qui se croisent, chacun d'entre eux voyant passer l'autre à une vitesse de 280 000 km/s. Imaginons que les deux vaisseaux se croisent de nouveau au bout d'une heure, toujours à la même vitesse ; chacun des cosmonautes doit alors voir l'horloge de l'autre retarder d'une demi-heure. Mais il n'est pas possible que les deux horloges soient chacune en retard sur l'autre. Alors à quoi doit-on s'attendre ? Ainsi s'énonce ce que l'on appelle le *paradoxe des horloges*.

En fait, il n'y a pas de paradoxe ! Lorsque les deux vaisseaux se croisent, les cosmonautes des deux vaisseaux peuvent bien jurer leurs grands dieux qu'ils ont vu que l'horloge d'en face était en retard, cela n'a aucune importance ; car, si les deux vaisseaux se croisent réellement, cela veut dire qu'ils ne se rencontreront jamais de nouveau ; on ne pourra jamais rassembler les deux horloges en un même endroit pour comparer au même instant leurs indications, et le paradoxe ne se révélera jamais. De fait, la théorie de la relativité restreinte d'Einstein ne s'applique qu'au cas de mouvements uniformes.

Mais imaginons que les deux vaisseaux se trouvent de nouveau à un instant au même endroit. Cela n'est possible qu'en introduisant une condition supplémentaire dans l'expérience : l'un des vaisseaux, au moins, doit subir une accélération. Supposons que ce soit le vaisseau B qui subisse cette accélération : B, après avoir croisé A, ralentit, fait demi-tour et accélère pour rattraper A. Rien n'empêche B de se choisir lui-même comme système de référence. De son point de vue, il est immobile et c'est A qui ralentit, fait demi-tour et le rattrape à reculons.

Si les deux vaisseaux étaient seuls dans l'Univers, le paradoxe existerait alors bel et bien. Mais tel n'est pas le cas. Et de ce fait A et B ne jouent pas des rôles symétriques. Quand B accélère, il accélère par rapport à A évidemment, mais aussi par rapport au reste de l'Univers. Si donc B se choisit comme système de référence, il lui faut tenir compte de ce que non seulement A, mais également tout le reste de l'Univers, toutes les galaxies sans exception, sont en accélération par rapport à lui. En somme, il y a B d'un côté, et le reste de l'Univers de l'autre. Et dans ces conditions, c'est l'horloge de B, et pas celle de A, qui retarde d'une demi-heure quand les vaisseaux se retrouvent au même endroit.

Ce phénomène risque d'affecter les vols spatiaux. Pour des astronautes voyageant à une vitesse voisine de celle de la lumière, le temps s'écoule beaucoup plus lentement que pour nous. Ces astronautes pourraient très bien atteindre une destination lointaine et revenir en un temps qui leur semblerait avoir duré seulement quelques semaines, mais qui, sur Terre, correspondrait à des siècles. Si le temps se ralentit effectivement lorsqu'on est en mouvement, on peut alors envisager d'effectuer en l'espace d'une vie le voyage à destination des étoiles. Mais évidemment, il faudrait alors quitter pour toujours le monde actuel et abandonner l'espoir de jamais revoir les personnes que l'on a connues, celles de la même génération. Ce serait vraiment l'embarquement pour le futur !

LA GRAVITATION ET LA THÉORIE D'EINSTEIN

La théorie de la relativité restreinte ne traite pas du mouvement accéléré sous l'effet de la gravitation. Le cas de ce mouvement a été traité par Einstein, un peu plus tard, en 1915 ; sa nouvelle théorie, théorie de la relativité générale, donne de la gravitation une idée complètement renouvelée : la gravitation apparaît comme une propriété de l'espace, et non plus comme une force d'interaction entre corps. Du fait de la présence de corps doués de masse, l'espace est courbé et les corps glissent en quelque sorte le long des lignes de plus grande pente de cet espace incurvé. Aussi étranges que ces idées d'Einstein puissent paraître à première vue, elles ont néanmoins permis d'expliquer un certain nombre de phénomènes que la théorie newtonienne de la gravitation n'arrivait pas à expliquer.

1846, l'année de la découverte de la planète Neptune (voir à ce sujet le chapitre 3) marque sans doute le sommet de la gloire de la théorie de Newton. Plus rien, après un tel éclat, ne semblait pouvoir remettre en cause la théorie de Newton. Pourtant, il restait encore un mouvement planétaire non expliqué : le mouvement de la planète Mercure, au point où elle s'approche le plus du Soleil – à son *périhélie* comme on dit –, varie d'une révolution à l'autre ; le mouvement avance au fur et à mesure que se déroulent les révolutions de Mercure autour du Soleil. Les astronomes avaient réussi à rendre compte à peu près de cette irrégularité en faisant intervenir l'attraction subie par Mercure du fait des planètes voisines.

Certes, on avait pensé à un moment, aux tout débuts de l'ère newtonienne, que les diverses perturbations, provoquées par l'attraction des planètes entre elles, pourraient bien avoir comme conséquence de déstabiliser le mécanisme fragile du système solaire. Mais, dans les premières années du XIXe siècle, Laplace avait montré que le système solaire n'est pas si fragile qu'il y paraît : les perturbations sont elles-mêmes cycliques et les irrégularités qu'elles provoquent ne dépassent jamais une certaine valeur dans une direction donnée ; si bien que, sur une longue période de temps, le système solaire est stable. Les astronomes étaient donc réellement persuadés au début du XXe siècle que les perturbations dues à l'attraction des planètes entre elles devaient pouvoir rendre compte de toutes les irrégularités observées.

Malheureusement, cette hypothèse ne permettait pas de rendre compte de l'avance du périhélie de Mercure : une fois incorporées toutes les perturbations provenant des autres planètes, il restait encore une avance du périhélie de Mercure de 43 secondes d'arc par siècle complètement inexpliquée. Cette avance, découverte en 1845 par Le Verrier, est vraiment infime : en quatre mille ans, l'avance prise ne dépasse pas la taille de la Lune. Mais c'était assez pour plonger les astronomes dans l'embarras.

Le Verrier avait émis l'idée que l'avance en question était due à la présence d'une planète encore inconnue qui, selon lui, devait être toute petite et se situer au voisinage de Mercure. Pendant plusieurs dizaines d'années les astronomes avaient recherché cette petite planète, baptisée Vulcain. Plus d'une fois on avait cru l'avoir enfin découverte. Mais à chaque fois, on s'était aperçu qu'il n'en était rien, et finalement on en était venu à penser que Vulcain n'existait tout simplement pas.

Or, voilà que la théorie de la relativité générale élaborée par Einstein apportait une réponse au problème de l'avance du périhélie de Mercure. Einstein avait en effet montré que le périhélie de tous les mouvements de rotation doit être plus grand que celui prédit par la théorie de Newton. En appliquant ce résultat général au cas particulier de Mercure, Einstein avait retrouvé exactement l'avance observée depuis un siècle. Les planètes plus éloignées du Soleil que Mercure ont un mouvement de périhélie encore moins important. En 1960, on a mesuré l'avance du périhélie de Vénus ; on a trouvé 8 secondes d'arc par siècle, ce qui est exactement le chiffre prédit par Einstein.

Mais ce n'est pas tout. La théorie d'Einstein prédisait également l'existence de deux phénomènes inconnus et stupéfiants. Le premier d'entre eux a trait au ralentissement des vibrations des atomes placés dans un champ intense de gravitation ; ce ralentissement devait se manifester par un déplacement des raies spectrales desdits atomes vers le rouge (ce que l'on appelle le *décalage vers le rouge*). Cherchant où trouver un champ

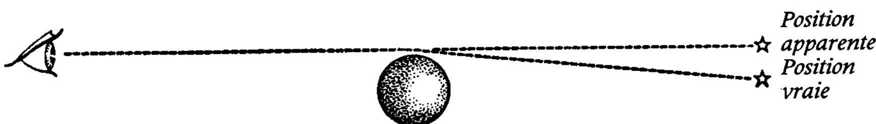
de gravitation suffisamment intense pour produire cet effet, Eddington suggéra de regarder du côté des naines blanches : la lumière quittant la surface de ces étoiles extrêmement condensées, où règne donc un champ gravitationnel extrêmement puissant, devait, si la théorie d'Einstein était juste, perdre une part importante de son énergie. En 1925, W.S. Adams, qui avait été le premier à montrer que ces étoiles ont une densité énorme, mit en évidence le décalage vers le rouge prévu par la théorie d'Einstein.

Quant à la seconde conséquence déroutante prévue par la théorie d'Einstein, elle donna lieu à une vérification des plus spectaculaires. Selon la théorie, une masse suffisamment importante devait provoquer une courbure des rayons lumineux passant en son voisinage, courbure s'élevant dans le cas du Soleil à 1,75 seconde d'arc (figure 8.4). Comment vérifier cette prédiction ? Il suffisait d'observer la lumière provenant d'étoiles situées derrière le Soleil pendant une éclipse de celui-ci ; en comparant les positions de ces étoiles telles qu'elles apparaîtraient alors à celles qu'on leur connaissait par ailleurs, on devait pouvoir mettre en évidence la courbure due au fait que la lumière provenant de ces étoiles aurait passé juste à raz de la surface du Soleil.

Einstein avait publié sa théorie de la relativité générale en 1915 ; il fallut attendre 1919 pour que soit effectuée cette vérification, juste après la fin de la Première Guerre mondiale. La British Royal Astronomical Society organisa une expédition chargée d'observer une éclipse totale de Soleil visible depuis l'île du Prince, possession portugaise située au large des côtes ouest-africaines. La position des étoiles s'avéra effectivement décalée. Einstein avait raison, une fois de plus.

Le même raisonnement que celui qui conduit à un déplacement de la position d'une étoile permet de dire que si une étoile se trouve derrière une autre, la lumière en provenance de l'étoile la plus éloignée devra contourner la plus proche, de telle sorte que l'étoile la plus éloignée paraîtra plus étendue qu'elle n'est en réalité. En d'autres termes, l'étoile la plus proche aura fait office de *lentille gravitationnelle*. Malheureusement, la taille apparente des étoiles est tellement petite que l'éclipse d'une étoile éloignée par une plus proche (vue de la Terre, évidemment) est un phénomène extrêmement rare. Mais la découverte des quasars a ouvert aux astronomes des possibilités nouvelles. Au début des années quatre-vingts, on a observé des quasars doubles constitués de parties ayant les mêmes caractéristiques. Il est raisonnable de penser qu'il s'agit en fait d'un seul et même quasar ; la lumière émise par ce quasar est déviée par une galaxie (ou un trou noir) située entre nous et le quasar, mais invisible par ailleurs ; l'image du quasar est alors double (de même qu'un mauvais miroir fournit une image dédoublée).

Figure 8.4. La courbure des rayons lumineux, prévue par la théorie de la relativité générale d'Einstein.



LES TESTS DE LA RELATIVITÉ GÉNÉRALE

Les premières vérifications de la théorie de la relativité générale portaient toutes sur des phénomènes astronomiques. Il fallut attendre longtemps avant que l'on soit en mesure de faire des vérifications semblables en laboratoire, dans des conditions telles que l'on puisse faire varier à volonté les paramètres. Ce n'est qu'en 1958 que l'on disposa de telles conditions, lorsque le physicien allemand Rudolf Ludwig Mössbauer démontra que, dans certains cas, un cristal peut émettre un faisceau de rayons gamma de longueur d'onde extrêmement bien définie. Normalement, lors de l'émission gamma, l'atome subit un recul et de ce fait, le spectre en longueur d'onde des rayons gamma est élargi. Mais, dans certaines conditions expérimentales, le cristal tout entier se comporte comme un seul atome, et comme le recul est alors réparti sur tous les atomes du cristal, il devient pratiquement nul ; si bien que le spectre des rayons gamma émis a une largeur pratiquement nulle. Par ailleurs, un faisceau possédant une telle monochromaticité est susceptible d'être absorbé, avec un taux d'absorption remarquablement élevé, par un autre cristal identique à celui qui lui a donné naissance. Si, pour une raison ou une autre, la longueur d'onde du faisceau incident est légèrement modifiée, alors l'absorption ne se produit pratiquement pas. Tel est ce que l'on appelle l'*effet Mössbauer.*

Pour un tel faisceau de rayons gamma émis vers le bas dans le champ de gravitation, la théorie de la relativité générale prévoit une augmentation d'énergie, et donc une diminution de la longueur d'onde. Sur une distance d'environ 300 mètres le faisceau doit voir son énergie augmenter suffisamment pour que, lors de son passage dans le second cristal, il ne soit pas absorbé.

Si maintenant, on entraîne le cristal émetteur vers le haut dans le champ de gravitation, sa longueur d'onde doit augmenter par effet Doppler. On peut alors, en ajustant la vitesse d'entraînement du cristal émetteur, arriver à comprendre l'effet de gravitation et, par conséquent, rétablir l'absorption par le second cristal.

Des expériences mettant à profit l'effet Mössbauer ont permis en 1960 de vérifier la théorie de la relativité générale d'Einstein avec une très grande précision. C'est certainement là la vérification la plus impressionnante jamais réalisée. Aussi Mössbauer s'est-il vu décerner le prix Nobel en 1961.

D'autres expériences, conduisant à des mesures extrêmement précises, confirment également la théorie de la relativité. Il s'agit entre autres de la mesure de la déviation d'un rayonnement radar au voisinage d'une planète ou de l'observation du comportement de pulsars binaires. Toutes ces mesures sont « limites » ; c'est pourquoi on a vu fleurir de nombreuses alternatives à la théorie d'Einstein. Malgré tout, la théorie d'Einstein reste, parmi toutes les théories existantes, celle qui est la plus simple du point de vue mathématique. Par ailleurs, chaque fois que l'on dispose de mesures qui permettent de discriminer entre deux théories, c'est toujours la théorie d'Einstein qui semble l'emporter. Trois quarts de siècle après sa formulation, la théorie de la relativité générale d'Einstein tient toujours la route, même si les physiciens continuent à la mettre en question (ce qui est normal). Attention : la théorie qui est périodiquement remise en question est la théorie de la relativité générale ; la théorie de la relativité

restreinte a été tellement souvent vérifiée que son exactitude aujourd'hui ne fait plus problème.

La chaleur

Jusqu'à présent, je n'ai pas tenu compte d'un phénomène qui pourtant, dans la vie courante, est indissociable de la lumière : la chaleur. Presque tous les objets lumineux, qu'il s'agisse d'une étoile ou d'une simple chandelle, émettent à la fois de la chaleur et de la lumière.

LA MESURE DES TEMPÉRATURES

Il a fallu attendre les temps modernes pour que la chaleur soit étudiée autrement que de façon qualitative. Jusque-là, on se contentait de dire : « C'est chaud », ou « C'est froid », ou bien « C'est plus chaud ». Pour passer au stade de l'étude quantitative, pour mesurer la température d'un corps, il a fallu d'abord trouver un phénomène correspondant à un changement qui semble varier de façon uniforme avec la température. C'est la dilatation (respectivement la contraction) que subit un corps quand on le chauffe (respectivement le refroidit) qui finalement a été adoptée.

C'est Galilée qui, le premier, a eu l'idée de repérer les changements de température en utilisant ce phénomène. En 1603, Galilée renversa un tube de verre contenant de l'air chauffé sur une cuve d'eau. En se refroidissant jusqu'à atteindre la température de la pièce, l'air se contractait et aspirait l'eau dans le tube. Galilée avait construit le premier *thermomètre* (littéralement « mesure de chaleur » en grec). Si la température de la pièce changeait, le niveau d'eau dans le tube renversé changeait aussi : quand la pièce se réchauffait, l'air dans le tube s'échauffait également et, repoussant l'eau dans la cuve, faisait que le niveau baissait ; si la pièce se refroidissait, l'air du tube se contractait et le niveau d'eau montait. Mais il y avait un problème : la cuve sur laquelle le tube de verre était retourné n'était pas fermée et restait en contact avec l'air de la pièce ; or, la pression de la pièce changeait ; ce qui modifiait le niveau de l'eau dans le tube indépendamment de l'effet dû à la température, et il était difficile de démêler les deux effets. A noter que le thermomètre de Galilée est le premier appareil scientifique en verre que l'on ait construit.

Dès 1654, le Grand Duc de Toscane, Ferdinand II, avait réalisé un thermomètre insensible aux variations de pression atmosphérique. Il s'agissait d'un ballon contenant une certaine quantité de liquide et qui était soudé à un tube en forme de tige. C'était alors la contraction ou la dilatation du liquide qui servait à repérer les variations de température. Mais les liquides changent moins de volume que les gaz pour une même différence de température. Pour remédier à cet inconvénient, le Grand Duc utilisait un ballon de bonne taille entièrement rempli de liquide ; ce ballon était rattaché à un tube de très faible section ; de la sorte, le liquide ne pouvait se dilater qu'en remplissant le tube fin, et les variations de niveau dans ce tube étaient parfaitement visibles, même si le liquide ne s'était que très peu dilaté.

A peu près à la même époque, le physicien anglais Robert Boyle avait réalisé un thermomètre très semblable, dont il se servait pour étudier la température du corps humain ; c'est lui qui le premier a montré que la température du corps humain est relativement constante, et nettement supérieure aux températures extérieures.

Aux tout débuts de la thermométrie, on utilisait comme liquides surtout l'eau et l'alcool. Mais l'eau se solidifie à une température relativement élevée et l'alcool se met à bouillir à une température assez basse. C'est pourquoi le physicien français Guillaume Amontons eut l'idée d'utiliser plutôt du mercure. Son thermomètre fonctionnait sur le même principe que celui de Galilée, à savoir que c'est la dilatation ou la contraction d'un certain volume d'air qui faisait varier le niveau du mercure dans le tube.

Et puis, en 1714, le physicien allemand Gabriel Daniel Fahrenheit combina les deux inventions, celle du Grand Duc et celle d'Amontons : un bulbe semblable à celui du thermomètre du Grand Duc était rempli de mercure, et c'est la dilatation ou la contraction du mercure lui-même qui modifiaient le niveau dans le tube fin. Fahrenheit eut aussi l'idée de placer une règle graduée le long de la tige de manière à pouvoir obtenir directement une lecture quantitative de la température.

Quant à savoir comment Fahrenheit s'y prit pour établir ce qui porte aujourd'hui le nom d'échelle Fahrenheit, la chose n'est pas claire et les historiens ne sont pas d'accord entre eux. Il est probable qu'il choisit comme zéro la température la plus basse qu'il pouvait atteindre avec les moyens de son laboratoire, c'est-à-dire celle obtenue en refroidissant un mélange d'eau et de sel. Il décida alors d'appeler 32 le point de congélation de l'eau pure et 212 son point d'ébullition. Cela présentait deux avantages ; d'abord entre les deux points particuliers relatifs à l'eau pure, il disposait de 180 graduations ; or 180, c'est aussi le nombre de degrés d'angle qui divisent l'angle plat ; ce qui semblait naturel, s'agissant de *degrés*. Ensuite, en procédant ainsi, la température du corps se trouvait être voisine de 100 (98,6 ° Fahrenheit très exactement).

La température du corps humain est remarquablement constante ; au point que l'on considère qu'un malade a de la fièvre dès que sa température dépasse d'un degré la température ordinaire. En 1858, le médecin allemand Karl August Wunderlich fut le premier à étudier l'évolution d'une maladie en relevant la température du patient. Dix ans plus tard, le médecin anglais Thomas Clifford Allbutt inventa le *thermomètre médical*, qui a ceci de particulier qu'il présente un resserrement très marqué entre le bulbe et la tige ; ainsi, le filet de mercure s'élève dans la tige lorsqu'on met le thermomètre en place et il ne peut redescendre lorsqu'on retire l'instrument du corps du patient ; le filet de mercure se sépare en deux au niveau du retrécissement et l'on peut alors lire aisément la température du malade. Aux États-Unis, l'échelle Fahrenheit est couramment utilisée pour mesurer les températures ; la température donnée par les bulletins météorologiques, en particulier, est en degrés Fahrenheit.

En 1742, l'astronome suédois Anders Celsius proposa d'utiliser une autre échelle. Le degré zéro est alors au point de congélation de la glace, et 100 représente le point d'ébullition de l'eau. De ce fait, le domaine de température dans lequel l'eau est liquide est divisé en cent ; c'est pourquoi cette échelle porte le nom d'*échelle centigrade* (ou *degrés centigrades*, du latin « cent marches ») (voir figure 6.4). Généralement on parle de degrés centigrades, mais la conférence internationale réunie en 1948 a décidé

de rebaptiser cette échelle et de l'appeler échelle Celsius (ou degrés Celsius) en l'honneur de son inventeur, de manière à établir un parallèle avec l'échelle Fahrenheit. Officiellement donc, il faudrait parler de degrés Celsius. De toute façon, l'abréviation symbolique reste la même : *C*. C'est cette échelle qui est la plus couramment utilisée dans le monde civilisé, et même les États-Unis s'y sont ralliés ou du moins essaient de s'y rallier. Pour les physiciens, il ne fait guère de doute que l'échelle Celsius est beaucoup plus agréable à manipuler que l'autre.

LES DEUX THÉORIES DE LA CHALEUR

La température donne une mesure de l'intensité de la chaleur mais pas de sa quantité. La chaleur s'« écoule » toujours des hautes températures vers les basses températures, de manière à réaliser l'égalité des températures, de la même façon que l'eau s'écoule toujours vers l'endroit où le niveau est le plus bas, de manière à réaliser l'égalité des niveaux d'eau. La chaleur se comporte toujours ainsi quelle que soit par ailleurs la quantité de chaleur contenue dans les corps considérés. Bien qu'une baignoire pleine d'eau tiède ait un contenu calorifique bien supérieur à celui d'une allumette enflammée, la chaleur s'écoule toujours de l'allumette vers la baignoire quand on approche l'allumette de la baignoire, et jamais en sens inverse.

C'est à Joseph Black, célèbre par ailleurs pour ses travaux sur les corps gazeux (voir le chapitre 5) que revient le mérite d'avoir le premier établi une distinction entre chaleur et température. Il établit en 1760 le fait que des corps de substances diverses voyaient leur température s'élever diversement lorsqu'on leur communiquait une même quantité de chaleur. Il faut fournir trois fois plus de chaleur à un gramme de fer pour élever sa température d'un degré Celsius, trois fois plus qu'il ne faut en fournir à un gramme de plomb pour élever sa température de la même façon. Quant au béryllium, il faut lui fournir pour le même résutlat trois fois plus de chaleur qu'au fer.

Par ailleurs, Black établit qu'il était possible de fournir de la chaleur à un corps sans pour autant élever sa température. Par exemple, lorsqu'on fait fondre de la glace, un apport supplémentaire de chaleur active le processus de liquéfaction, mais n'augmente pas la température de la glace. Au bout du compte, la glace se retrouve entièrement transformée en eau ; mais pendant tout le temps que dure la liquéfaction, la température de la glace ne dépasse jamais 0 °C. De même pour l'eau en train de se vaporiser : si on lui fournit de la chaleur, la quantité d'eau qui passe à l'état de vapeur augmente, mais la température du liquide reste constante pendant toute l'opération de vaporisation.

Le développement des machines à vapeur (voir chapitre 9), qui date à peu près de la même époque que les expériences de Black, joua un grand rôle dans la prise de conscience par les savants de l'époque de l'importance des problèmes liés à la chaleur et à la température. De là datent les premières spéculations scientifiques sur la nature de la chaleur, spéculations qui ne sont pas sans rappeler celles qui en d'autres temps avaient porté sur la nature de la lumière.

Dans un cas comme dans l'autre, on vit se développer deux théories rivales. Selon la première théorie de la chaleur, celle-ci devait être

considérée comme une substance matérielle susceptible d'être ajoutée ou enlevée, transférée d'un corps à un autre. C'est la théorie dite du *calorique*, d'après le mot qui en latin signifie « chaleur ». Selon les tenants de cette théorie, lorsque du bois brûle, par exemple, une certaine quantité de calorique passe du bois dans la flamme, et de là dans la bouilloire placée au-dessus de la flamme, puis finalement de la bouilloire dans l'eau qu'elle contient. Au fur et à mesure que le contenu de l'eau en calorique augmente, l'eau se transforme en vapeur.

Quant à la deuxième théorie, elle a pour origine deux expériences célèbres, réalisées dans le courant du XVIII[e] siècle : il s'agit de la théorie *cinétique* de la chaleur, qui considère que la chaleur est l'effet de vibrations. La première de ces expériences a été réalisée par le physicien et aventurier Benjamin Thompson, d'origine américaine, qui avait quitté le pays au moment de la Révolution et s'était fait donner le titre de Comte Rumford, avant de parcourir toute l'Europe à la recherche de son destin. Alors qu'il dirigeait la fabrication de canons en Bavière, en 1798, il constata que le forage des canons s'accompagnait d'un dégagement de chaleur, suffisant pour porter à ébullition dix litres d'eau en moins de trois heures. D'où pouvait bien venir tout ce calorique ? se demanda Rumford. Après avoir réfléchi, il conclut que la chaleur devait être une forme de vibration engendrée et intensifiée par le frottement mécanique de l'appareil à forer contre la paroi du canon.

L'année suivante, le chimiste Humphry Davy réalisa une expérience encore plus probante. Davy disposait de deux morceaux de glace portés à une température inférieure à la température de fusion de la glace ; il les faisait frotter l'un contre l'autre, non pas à la main, mais à l'aide d'un dispositif mécanique, de façon qu'aucun « calorique » ne puisse s'échapper et se communiquer à la glace. Il constata alors que la glace fondait, sous l'effet de ce seul frottement. Il en conclut, lui aussi, que la chaleur devait être une forme de vibration, et pas une substance. De fait, en dépit des résultats incontestables de cette expérience, la théorie du calorique résista pendant longtemps et se maintint jusque dans le milieu du XIX[e] siècle.

LA CHALEUR COMME FORME D'ÉNERGIE

En dépit de ce contresens sur la nature de la chaleur, les physiciens, pendant tout le XIX[e] siècle, apprirent un certain nombre de choses concernant la chaleur ; de la même façon que leurs prédécesseurs avaient découvert un certain nombre de propriétés intéressantes concernant la lumière (les lois de la réflexion et celles de la réfraction, par exemple), tout en se trompant sur la véritable nature de la lumière. Les deux savants français Jean Baptiste Joseph Fourier en 1822, et Nicolas Léonard Sadi Carnot en 1824, étudièrent la manière dont s'écoule la chaleur et firent deux découvertes importantes. De fait, on considère généralement que Sadi Carnot est le père de la *thermodynamique* (ainsi nomme-t-on la science de la chaleur, d'après une expression grecque qui signifie « mouvement de la chaleur »). C'est lui qui a donné à l'étude des machines à vapeur ses assises théoriques.

Dans le courant des années 1840, les physiciens se trouvèrent confrontés au problème de savoir comment s'effectuait la transformation de la chaleur fournie à la vapeur en travail mécanique, capable de faire se déplacer

le piston d'une machine. Y avait-il une limite supérieure à la quantité de travail que l'on pouvait obtenir à partir d'une quantité donnée de chaleur ? Quant au processus inverse, comment s'effectuait-il ? Comment s'effectuait la conversion de travail en chaleur ?

Joule passa trente-cinq ans de sa vie à essayer de transformer diverses formes d'énergie en chaleur, refaisant avec soin les expériences un peu grossières de Rumford. Il mesura la quantité de chaleur produite par le passage d'un courant électrique. Il fit chauffer de l'eau et du mercure en les agitant avec des roues à aubes, ou en obligeant l'eau à s'écouler dans un tube étroit. Il chauffa de l'air en le comprimant, etc., etc. Dans chaque cas, il calcula le travail mécanique fourni au système et la quantité de chaleur obtenue. Il découvrit alors qu'une certaine quantité de travail, quelle qu'en soit la nature, fournit toujours la même quantité de chaleur. Joule venait de déterminer ce que nous appelons *l'équivalent mécanique de la chaleur*.

Puisque la chaleur pouvait être convertie en travail, il fallait bien la considérer comme une forme d'énergie (*énergie* vient d'un mot grec qui signifie « qui contient du travail »). L'électricité, le magnétisme, la lumière, le mouvement, qui tous sont capables de fournir du travail, devaient aussi être considérés comme des formes d'énergie. Quant au travail, puisqu'il est possible de le transformer en énergie, il constitue également une forme d'énergie.

Ces considérations redonnaient vie à une idée dont on soupçonnait bien depuis Newton qu'elle devait être juste, à savoir l'idée selon laquelle l'énergie est *conservée*, c'est-à-dire ne peut être ni créée ni détruite. Ainsi, par exemple, un corps en mouvement a de *l'énergie cinétique* (littéralement, « énergie de mouvement »), selon une terminologie introduite en 1856 par Lord Kelvin. Un corps qui s'élève dans le champ de la pesanteur voit son mouvement ralentir et son énergie cinétique s'évanouir progressivement ; mais, en même temps que son énergie cinétique diminue, il gagne de l'énergie de position ; on veut dire par là que sa position étant élevée par rapport à la surface de la Terre, il peut retomber et par là même récupérer son énergie cinétique. En 1853, le physicien écossais William John Macquorn Rankine proposa d'appeler cette énergie de position *énergie potentielle*. Pendant un temps, il put paraître que la somme énergie cinétique + énergie potentielle (ce que l'on appelle *l'énergie mécanique* d'un corps) était conservée, et on put parler de la loi de conservation de l'énergie mécanique. En réalité, l'énergie mécanique n'est qu'approximativement conservée, car une fraction de cette énergie sert toujours à vaincre les frottements ou la résistance de l'air.

L'expérience de Joule intervenant dans ce contexte montrait clairement que l'énergie est effectivement conservée dès lors qu'on prend en compte cette forme particulière d'énergie qu'est la chaleur ; en effet, tous les cas où de l'énergie mécanique n'est pas conservée sont liés à l'apparition d'un dégagement de chaleur, que ce soit du fait des frottements ou du fait de la résistance de l'air. Si l'on tient compte de cette chaleur, il apparaît que l'énergie est bien conservée : elle ne peut être ni créée, ni détruite. C'est le physicien Julius Robert von Mayer qui exprima le premier cette idée en termes clairs, en 1842. Malheureusement, Mayer n'avait guère d'expérience à présenter à l'appui de son idée et il n'avait que peu de soutien parmi les universitaires de son temps (Joule lui-même, qui n'avait pas fait une carrière universitaire, mais était brasseur de profession, eut beaucoup de mal à se faire publier).

Il fallut attendre 1847 pour qu'un universitaire suffisamment respectable énonce cette même idée. Cet universitaire se nomme Heinrich von Helmholtz et la loi qui porte son nom, la *loi de conservation de l'énergie,* s'énonce ainsi : chaque fois qu'une certaine quantité de travail disparaît en un endroit, une quantité égale d'énergie apparaît ailleurs. Cette loi porte aussi le nom de *premier principe de la thermodynamique*. C'est l'une des pierres angulaires de l'édifice théorique de la physique que ni la relativité ni la théorie quantique n'ont réussi à ébranler.

Cependant, s'il est vrai que n'importe quelle quantité de travail peut être intégralement transformée en chaleur, l'inverse, la conversion totale de chaleur en travail, n'est pas vrai. Lorsque de la chaleur est convertie en travail, une partie de la quantité de chaleur ne peut être utilisée et elle est tout simplement perdue. Dans le cas d'une machine à vapeur, la conversion de chaleur en travail ne s'effectue que jusqu'à ce que la température ait atteint la température du milieu environnant ; arrivée à ce point-là, la conversion s'arrête, et ce, bien qu'il reste encore beaucoup de chaleur à convertir. Et même à l'intérieur du domaine de températures qui permet d'extraire du travail de la vapeur chauffée, une partie de la chaleur ne se transforme pas en travail, mais sert simplement à chauffer l'air environnant et la machine elle-même, et à vaincre les frottements qui se développent au niveau du cylindre et du piston.

Dans tout processus de conversion d'une forme d'énergie en une autre, qu'il s'agisse de convertir de l'énergie électrique en énergie lumineuse ou de l'énergie magnétique en énergie cinétique, une fraction de l'énergie n'est pas convertie. Cette fraction n'est cependant pas perdue (ce serait contraire au premier principe de la thermodynamique) ; elle est simplement convertie en chaleur, c'est-à-dire dissipée par le milieu environnant.

La capacité d'un système à fournir du travail est ce que l'on appelle son *énergie libre*. La fraction d'énergie qui est inévitablement dissipée sous forme de chaleur non utilisable se retrouve dans ce que l'on appelle *l'entropie,* selon une désignation introduite pour la première fois par le physicien allemand Rudolf Julius Emmanuel Clausius.

Clausius a montré que dans tout processus mettant en œuvre un flux d'énergie, il existe toujours des pertes ; de ce fait, l'entropie de l'Univers ne cesse d'augmenter. Telle est l'essence du *second principe de la thermodynamique* ; on parle également à ce propos de « mort thermique de l'Univers ». Heureusement, cependant, la quantité d'énergie utilisable (essentiellement fournie par les étoiles) est telle que nous en aurons encore suffisamment pendant quelques milliards de milliards d'années.

LA CHALEUR ET LE MOUVEMENT DES MOLÉCULES

Ce n'est qu'avec la théorie atomique de la matière que se dégagea une réelle compréhension de la nature de la chaleur. Dans la conception atomique, en effet, les molécules qui constituent un gaz sont animées d'un mouvement incessant, toujours en train de se heurter les unes les autres ou de rebondir sur les parois du récipient dans lequel se trouve enfermé le gaz en question. Cette conception n'était pourtant pas nouvelle : déjà en 1738, le mathématicien suisse Daniel Bernouilli avait expliqué les propriétés des gaz par de telles considérations. Mais Bernouilli était en avance sur son temps, et il fallut attendre le milieu du XIXe siècle pour

que soit élaborée par Maxwell et Boltzmann une théorie mathématique des gaz, théorie qui porte le nom de *théorie cinétique des gaz* (cinétique, pour rappeler qu'il s'agit du mouvement des molécules). Des résultats de la théorie cinétique des gaz, il découlait clairement que ce que nous appelons chaleur n'est autre que le mouvement des molécules. La théorie du calorique venait de recevoir un coup fatal et très vite, il devint clair pour tout le monde que la chaleur n'est qu'une forme de vibration : agitation désordonnée des molécules dans les gaz et les liquides, mouvement plus ordonné d'aller et retour dans le cas des solides.

Quand on chauffe un solide, arrive un moment où ce mouvement d'aller et retour des atomes autour d'une position moyenne devient suffisamment important pour que les liaisons entre atomes soient rompues ; le solide se met à fondre et à devenir liquide. Plus les atomes du solide sont solidement liés entre eux, plus ce solide est difficile à fondre, plus son point de fusion est élevé.

Dans un liquide, les molécules peuvent se déplacer librement les unes par rapport aux autres en glissant les unes sur les autres. Quand on chauffe un liquide, arrive un moment où les mouvements moléculaires ainsi provoqués deviennent assez forts pour libérer les molécules de leurs attaches à l'ensemble du liquide ; elles en sortent, le liquide bout. Ici encore, le point d'ébullition d'un liquide est d'autant plus élevé que les forces qui lient ses molécules entre elles sont grandes.

Lors du changement de l'état solide à l'état liquide, toute l'énergie fournie par la chaleur sert à rompre les liaisons intermoléculaires ; c'est pourquoi la chaleur fournie à la glace pour la faire fondre ne fait pas s'élever la température de la glace. Il en va de même dans le cas de l'évaporation.

Nous sommes maintenant en mesure de préciser la différence entre température et chaleur. La chaleur n'est autre que l'énergie de mouvement de l'ensemble des molécules du corps considéré. La température, elle, représente l'énergie moyenne d'une molécule de ce corps. Ainsi, par exemple, deux litres d'eau portés à une température de 60 °C renferment deux fois plus de chaleur qu'un litre d'eau porté à la même température de 60 °C (en effet, il y a dans le premier cas deux fois plus de molécules en mouvement) ; mais les deux litres et le litre sont tous les deux à la même température ; cela veut dire que l'énergie moyenne d'une molécule est la même dans les deux cas.

Tout composé chimique renferme de l'énergie du fait de sa structure de composé : il s'agit de l'énergie associée aux forces qui lient les atomes du composé entre eux. Si ces liaisons entre atomes sont modifiées, et si ce réarrangement correspond à un abaissement de l'énergie, le reste de l'énergie se manifestera sous forme soit de lumière, soit de chaleur, soit même sous les deux formes à la fois. Il arrive que cette énergie soit dégagée de façon suffisamment brutale pour provoquer une explosion.

On est capable de calculer l'énergie chimique renfermée par n'importe quelle substance chimique et d'en déduire l'énergie qui sera dégagée dans telle ou telle réaction. Ainsi par exemple, la combustion du bois met en jeu la rupture des liaisons entre les atomes de carbone du charbon et des liaisons entre les atomes d'oxygène de l'air avec lesquels les carbones se combinent. Or, l'énergie des liaisons dans le composé ainsi produit (le dioxyde de carbone) est inférieure à celle des liaisons dans l'état initial. C'est cette différence qui apparaît sous forme de chaleur et de lumière.

En 1876, le physicien américain Josiah Willard Gibbs a développé une théorie de la *thermodynamique chimique* avec tant de détails et de précision que cette discipline est passée d'un seul coup de l'état de non-existence à celui de théorie opérante. L'article assez long dans lequel Gibbs expose ses travaux représentait une telle avance par rapport à l'état de cette science à l'époque que Gibbs eut toutes les peines du monde à le faire publier par l'académie des arts et des sciences du Connecticut. Même après la parution de l'article, la nature des arguments mathématiques et le caractère réservé de Gibbs firent que le sujet resta non exploité pendant longtemps, jusqu'à ce que Ostwald découvre en 1883 l'existence de l'article de Gibbs et en fasse paraître une traduction en allemand ; ce qui proclama aux yeux du monde entier l'importance des travaux de Gibbs.

Nous ne citerons qu'un seul exemple montrant l'importance du travail de Gibbs. Grâce à ses équations, Gibbs avait été capable d'établir de façon rigoureuse les lois qui régissent l'équilibre entre diverses substances coexistant dans plusieurs phases (c'est-à-dire, par exemple, sous forme solide et en solution, ou bien sous forme de deux liquides non miscibles, etc.). Cette *règle des phases* inventée par Gibbs est ce qui permet à la métallurgie et à bien des branches de l'industrie chimique d'exister.

Masse et énergie

Avec la découverte de la radioactivité en 1896 (voir chapitre 6), des questions d'un ordre totalement nouveau se posèrent aux physiciens. Les substances radioactives, l'uranium et le thorium, libéraient des particules possédant une énergie étonnamment élevée. De plus, Marie Curie avait découvert que le radium émettait de façon permanente de grandes quantités de chaleur : 30 grammes de radium libéraient environ 4 000 calories par heure, et cela pendant des heures et des heures, des semaines, des années même. La réaction chimique la plus énergétique n'aurait pu libérer qu'une millionième partie de l'énergie libérée par le radium. Fallait-il voir là l'indice que la loi de conservation de l'énergie était violée ?

Non moins surprenant était le fait que cette production d'énergie, contrairement à ce qui se passait dans le cas des réactions chimiques, ne dépendait pas de la température ; elle se poursuivait, égale à elle-même, aussi bien à température ordinaire qu'à la température de l'hydrogène liquide !

Visiblement, on avait affaire là à un autre type d'énergie, complètement différente de l'énergie chimique. Fort heureusement, les physiciens n'eurent pas à attendre longtemps la solution de cette énigme. Une fois de plus, c'est Einstein qui trouva la solution, en tirant les conséquences de sa théorie de la relativité restreinte. Einstein démontra de façon mathématique que la masse peut être considérée comme une forme d'énergie, une de plus, très fortement concentrée, puisqu'une très petite quantité de masse est susceptible d'être transformée en une quantité énorme d'énergie.

La relation qui, selon Einstein, lie la masse à l'énergie est certainement l'équation la plus connue de la physique :

$$E = mc^2$$

où E représente l'énergie (en ergs) et m la masse (en grammes) ; c représente la vitesse de la lumière (en centimètres par seconde).* Si l'on utilise un autre système d'unités, la relation reste la même.

Etant donné que la vitesse de la lumière est de 300 000 km/s, le facteur de conversion, pour un gramme de matière, a une valeur de 9×10^{20} ergs. Certes, l'*erg* est une unité toute petite ; mais l'énergie contenue dans un gramme de matière est capable de maintenir allumée pendant 2 850 ans une lampe électrique de 1 000 watts; cela donne une idée plus familière de l'énormité du facteur de conversion. On pourrait aussi dire que si l'on arrivait à convertir complètement l'énergie contenue dans un gramme de matière, cela fournirait la même énergie que 2 000 tonnes de pétrole.

L'équation d'Einstein mettait à bas l'une des lois de conservation les plus intouchables de la physique, à savoir la loi de conservation de la masse énoncée par Lavoisier, selon laquelle rien ne se perd, rien ne se crée. Pourtant, à y regarder de près, toute réaction chimique dégageant de l'énergie change une petite fraction de la masse des réactifs en énergie ; si l'on était capable de peser avec assez de précision les produits de la réaction, on verrait que leur masse n'est pas égale à celle des produits d'origine. Mais la masse ainsi perdue lors d'une réaction chimique est si faible qu'aucune des techniques expérimentales dont on disposait au XIXe siècle n'aurait permis de la mettre en évidence. Avec la découverte d'Einstein, on avait affaire à un phénomène d'un tout autre ordre, portant non plus sur des réactions chimiques mais sur des réactions nucléaires, pour lesquelles la perte d'énergie est telle qu'elle en devient perceptible.

En posant l'équivalence masse-énergie, Einstein avait donc réussi à combiner en une seule loi de conservation – la *loi de conservation de la masse-énergie* – ce qui jusqu'alors avait pu passer pour deux lois de conservation distinctes. Non seulement le premier principe de la thermodynamique restait encore vrai, mais il s'en trouvait même renforcé.

La conversion de la masse en énergie reçut une confirmation expérimentale en 1925 grâce aux expériences effectuées par Aston à l'aide de son spectrographe de masse : Aston était en mesure de déterminer avec une extrême précision la masse des noyaux atomiques à partir de la déviation de ces mêmes noyaux dans un champ magnétique ; il montra ainsi que les divers noyaux ne sont pas des multiples exacts des masses du neutron et du proton qui les composent.

Arrêtons-nous un instant sur cette question des masses respectives du neutron et du proton. Pendant un siècle, on avait mesuré la masse des atomes et des particules subatomiques en supposant que la masse de l'atome d'oxygène avait exactement la valeur 16,00000... (voir le chapitre 6). En 1929, cependant, Giauque montra que l'oxygène est constitué de trois isotopes – l'oxygène 16, l'oxygène 17 et l'oxygène 18

* En France, on utilise le Système International d'Unités (SI) ; alors, E est en joules, m en kilogrammes et c en mètres par seconde. (*N.D.T.*)

– et que le poids atomique de l'oxygène est la moyenne pondérée des nombres de masse de ces trois isotopes.

Certes, l'oxygène 16 est de loin l'isotope prédominant, constituant à lui seul 99,759 % de tous les atomes d'oxygène ; en sorte que si l'oxygène naturel se voit attribuer le poids atomique de 16,00000..., l'isotope 16 a nécessairement un poids atomique très voisin de 16 (à cause de l'infime fraction d'oxygène 17 et 18 contenus dans l'oxygène naturel, l'isotope 16 a un poids atomique très légèrement supérieur à 16). C'est la raison pour laquelle, pendant au moins une génération, après la découverte des isotopes de l'oxygène, les chimistes ne se sont pas inquiétés et ont continué à fonctionner sur l'ancienne base des poids atomiques, ce que l'on appelait alors les poids atomiques chimiques.

Les physiciens, eux par contre, eurent une tout autre réaction. Il leur sembla préférable de fixer à 16 exactement le poids atomique de l'isotope 16 de l'oxygène et d'en déduire les autres poids atomiques (que l'on a appelés poids atomiques physiques). Sur cette base, le poids atomique de l'oxygène naturel, à cause de ses traces d'isotopes plus lourds, est de 16,0044. De façon générale, les poids atomiques physiques sont 0,027 % plus grands que les poids atomiques chimiques.

En 1961, physiciens et chimistes ont signé un compromis : les poids atomiques ont été dès lors déterminés en prenant 12,00000 pour l'isotope 12 du carbone. Ce compromis présente l'avantage, moyennant un changement d'élément de référence, de garder aux anciens poids atomiques à peu près la même valeur qu'auparavant ; le poids atomique de l'oxygène naturel, par exemple, est maintenant de 15,9994.

Ceci dit, considérons pour commencer un atome de l'isotope 12 du carbone, de poids atomique exactement égal à 12, donc. Son noyau comporte six neutrons et six protons. Les mesures effectuées au spectrographe de masse permettent d'établir que si la masse de l'isotope 12 du carbone vaut exactement 12, alors la masse du proton vaut 1,007825 et celle du neutron 1,008665. Six protons ont donc une masse de 6,046950 et six neutrons une masse de 6,051990 ; soit au total 12,104940 pour le carbone 12... dont la masse a été fixée à 12 exactement. Alors ? Où sont passés les 0,104940 qui manquent ?

Cette masse manquante est ce que l'on appelle le *défaut de masse*. Le défaut de masse divisé par le nombre de masse d'un noyau donne le défaut de masse par nucléon de ce noyau, ce que l'on appelle aussi son *facteur d'entassement*. La masse n'a évidemment pas disparu ; elle a simplement été convertie en énergie, selon la loi énoncée par Einstein ; si bien que le défaut de masse mesure en réalité *l'énergie de liaison* de ce noyau, c'est-à-dire l'énergie qu'il faudrait fournir au noyau pour le dissocier en ses protons et neutrons constituants (en effet, on créerait alors une quantité de masse correspondant à cette énergie).

Aston détermina les défauts de masse par nucléon caractéristiques d'un certain nombre de noyaux ; il s'aperçut que cette énergie commençait d'abord par croître assez rapidement à partir de l'hydrogène pour atteindre un maximum au niveau du fer et redescendre ensuite, plus lentement, du côté des numéros élevés du tableau périodique. Ce qui prouve que la conversion masse-énergie est susceptible de fournir de l'énergie pour des éléments situés aux deux bouts du tableau périodique.

Considérons par exemple le cas de l'uranium 238. Ce noyau peut se scinder en plusieurs étapes jusqu'à aboutir au plomb 206. Ce faisant, il

émet huit particules alpha (il y a aussi émission de particules bêta, mais celles-ci sont si légères qu'elles peuvent être, ici pour ce qui nous intéresse, simplement ignorées). La masse du plomb 206 est 205,9745, et celle des huit particules alpha 32,0208 ; soit au total 237,9953. Mais la masse de l'uranium 238 dont ces diverses particules sont issues est 238,0506 ; soit un défaut de 0,0553, lequel représente exactement l'énergie libérée par la fission de l'uranium.

Lorsque l'uranium subit d'autres désintégrations conduisant à des atomes encore plus petits que le plomb (ce qui est le cas dans le procédé de fission), l'énergie libérée est évidemment plus importante. De même, lorsque l'hydrogène est transformé en hélium, ce qui se produit à chaque instant dans le cœur des étoiles, le défaut de masse est encore plus grand et l'énergie libérée encore plus importante.

Les physiciens sont maintenant entraînés à vérifier à chaque instant la loi de conservation de la masse-énergie. C'est ainsi, par exemple, que lors de la découverte du proton en 1934, on vérifia que son annihilation par un électron avec production de rayons gamma correspondait bien à une énergie, pour ces rayons gamma, égale à la masse des deux particules. A l'inverse, comme le souligna Blackett, de la masse peut être créée à partir de l'énergie. Des rayons gamma d'énergie convenable peuvent, dans certaines circonstances, disparaître complètement en donnant naissance à une paire *électron-positron*. Des énergies plus élevées, telles que celles que l'on trouve dans les rayons cosmiques, ou telles que celles qui sont communiquées à des particules circulant dans un synchrotron (voir chapitre 7), sont capables d'engendrer des particules de masse plus élevée, des mésons ou des antiprotons, par exemple.

Rien d'étonnant, dans ces conditions, à ce que lorsqu'on se trouva confronté à une situation où la loi de conservation de la masse-énergie n'était pas respectée (comme ce fut le cas lorsqu'on examina l'émission bêta), on se soit vu obligé d'inventer une particule destinée à assurer cette fameuse conservation (voir chapitre 7).

S'il était besoin d'une preuve supplémentaire de la réalité du phénomène de conversion masse-énergie, la bombe atomique nous fournirait un argument, dont le moins qu'on puisse en dire est qu'il est sans réplique.

Ondes et particules

Dans le courant des années vingt, la physique était entièrement dominée par l'idée de dualité. Planck avait montré que la lumière se comporte à la fois comme une onde et comme une particule. Einstein avait montré que l'énergie et la masse ne sont que l'endroit et l'envers d'une même médaille, et que l'espace et le temps ne peuvent être dissociés. Du coup, les physiciens étaient tentés de voir des dualités un peu partout.

En 1923, le physicien français Louis Victor de Broglie démontra que, de la même façon que la lumière peut dans certains cas présenter des caractéristiques corpusculaires, les particules de matière, des électrons par exemple, peuvent présenter des caractéristiques ondulatoires. De Broglie prévoyait même que la longueur d'onde des ondes associées aux particules devait être inversement proportionnelle au produit de leur masse par leur

vitesse (autrement dit, à leur quantité de mouvement). De Broglie calcula la longueur d'onde associée à des électrons animés d'une vitesse « raisonnable » ; elle devait être du même ordre de grandeur que celle des rayons X.

Cette prédiction, pour le moins étonnante, fut confirmée en 1927. Clinton Joseph Davisson et Lester Halbert Germer, tous deux des Laboratoires Bell, étaient en train de réaliser une expérience de bombardement du nickel par des électrons lorsqu'à la suite d'un accident, ils durent chauffer l'échantillon de nickel métallique. L'échantillon, une fois chauffé, se présenta sous forme de monocristal.

Ce monocristal de grande taille avait les caractéristiques requises pour une expérience de diffraction d'électrons ; en effet dans un cristal, les distances interatomiques sont de l'ordre de grandeur des longueurs d'ondes associées à des électrons de vitesse modérée. C'est ce que confirmait l'expérience : les électrons ne traversaient pas le cristal comme l'auraient fait des particules ; ils se comportaient comme des ondes, et le film placé derrière l'échantillon portait la marque de figures d'interférence caractéristiques, avec des alternances de bandes noires et de bandes brillantes, tout comme si le cristal avait été soumis à des rayons X.

Or, l'existence de telles figures d'interférence était ce qui, un siècle auparavant, avait permis à Young de mettre en évidence la nature ondulatoire de la lumière. Dans le cas présent, l'existence de ces mêmes interférences apportait la preuve de la nature ondulatoire des électrons. La mesure de la distance entre deux bandes noires successives permettait de calculer la longueur d'onde associée aux électrons : Davisson et Germer trouvèrent une valeur de 1,65 angström, coïncidant presque avec la valeur prévue par de Broglie.

La même année, le physicien britannique George Paget Thomson démontra également, de façon totalement indépendante et par des méthodes différentes, la nature ondulatoire des électrons. De Broglie reçut le prix Nobel en 1929 ; quant à Davisson et Germer, ils partagèrent cette même récompense en 1937.

LA MICROSCOPIE ÉLECTRONIQUE

La découverte, tout à fait inattendue, d'un nouveau dualisme fut immédiatement mise à profit pour la réalisation d'un nouveau type de microscope : le microscope électronique. Les microscopes optiques, ordinaires, ne sont d'aucun secours (je l'ai déjà indiqué) dès que l'on cherche à observer des objets d'une taille inférieure à celle que les ondes lumineuses permettent de délimiter avec acuité. Plus on diminue la taille des objets à observer, plus ceux-ci deviennent flous, car la lumière les contourne. C'est ce qu'a établi en 1878 le physicien allemand Ernst Karl Abbe. Pour pallier cette difficulté, il faut évidemment avoir recours à des rayonnements de longueur d'onde de plus en plus courte ; on obtient ainsi un pouvoir de résolution supérieur. Dans un microscope ordinaire, on peut distinguer deux points séparés par 2/10 000 de millimètre. Avec un microscope fonctionnant dans l'ultraviolet, on abaisse cette limite à 1/10 000. Évidemment, avec des rayons X, on pourrait réaliser de meilleures performances ; malheureusement on ne sait pas construire de lentilles fonctionnant dans ce domaine. Mais on peut espérer augmenter

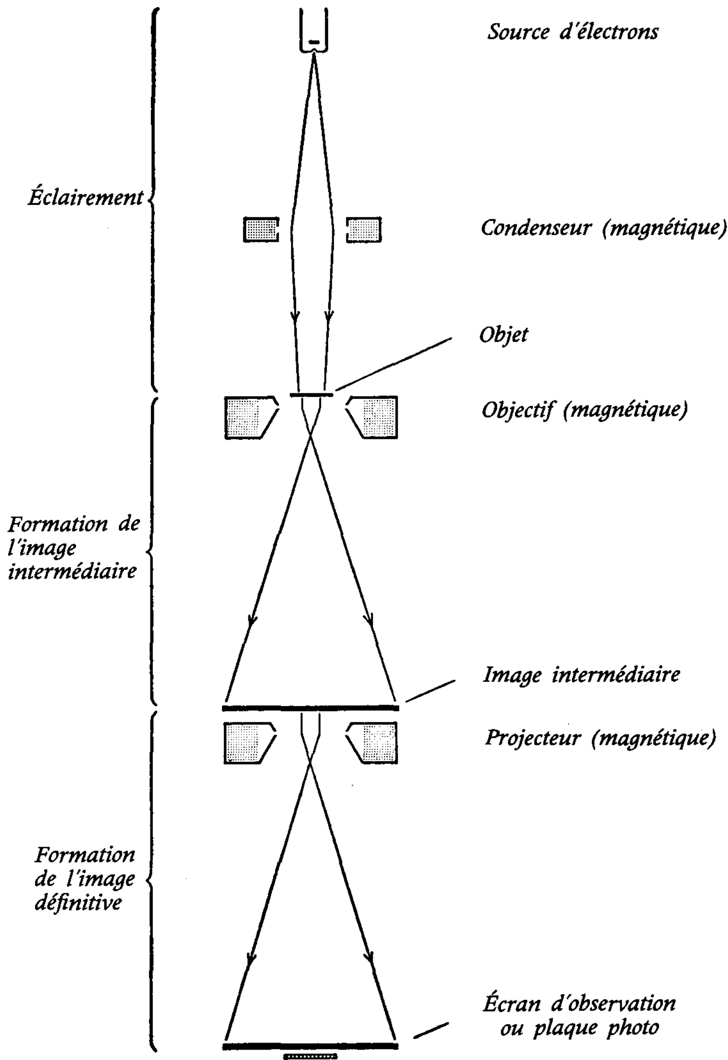


Figure 8.5. Schéma de fonctionnement d'un microscope électronique. Une première lentille magnétique transforme le faisceau divergent d'électrons en faisceau parallèle. L'objectif (magnétique) se comporte comme une lentille convergente et produit de l'objet une image agrandie, qui est ensuite reprise par un projecteur magnétique. L'image finale est projetée soit sur un écran fluorescent, soit sur une plaque photo.

la résolution en utilisant les ondes associées aux électrons ; celles-ci ont à peu près les mêmes longueurs d'onde que les rayons X et sont d'un maniement plus aisé. En particulier, on peut construire des lentilles magnétiques qui font se courber les *pinceaux d'électrons*, parce que les électrons sont des particules chargées.

De même que l'œil peut, à l'aide de rayons convenablement déviés par des lentilles, percevoir une image agrandie d'un objet, de même il est possible d'enregistrer sur une plaque photographique l'image d'un objet « éclairé » par des ondes électroniques convenablement déviées par des lentilles magnétiques. Et comme les longueurs d'onde des électrons peuvent être beaucoup plus petites que celles de la lumière ordinaire, la résolution obtenue dans un microscope électronique surpasse largement celle que l'on obtient avec un microscope ordinaire.

Le premier *microscope électronique*, de fabrication grossière, a été réalisé en Allemagne en 1932 par Ernst Ruska et Max Knoll ; cet instrument avait un grossissement de 400. Le premier microscope performant date de 1937 ; il fut construit par James Hillier et Albert F. Prebus de l'université de Toronto ; cet instrument avait un grossissement de 7 000, plus de trois fois supérieur à celui des microscopes ordinaires les plus performants. Très rapidement, dès 1939, les microscopes électroniques furent commercialisés, et Hillier et ses collègues en arrivèrent même à fabriquer des microscopes ayant un grossissement de 2 000 000.

Dans un microscope électronique ordinaire, les électrons sont focalisés sur l'échantillon qu'ils traversent. Il existe un autre type de microscope électronique où les électrons ne font que « balayer » l'échantillon, comme le font les électrons dans un poste de télévision. Cet instrument, qui porte le nom de *microscope à balayage*, repose sur une idée d'abord avancée par Knoll dès 1938, mais qui ne fut réalisée qu'en 1970 par le physicien américano-britannique Albert Victor Crew. L'observation par microscopie à balayage détériore moins l'objet que ne le fait la microscopie conventionnelle ; de plus, l'observation de l'objet dans l'espace y est meilleure qu'en microscopie conventionnelle. C'est avec de tels instruments que l'on a réussi à voir les atomes un à un.

L'ASPECT ONDULATOIRE DES ÉLECTRONS

On ne sera pas étonné d'apprendre que la dualité onde-corpuscule fonctionne dans les deux sens, c'est-à-dire que des phénomènes que l'on considère généralement comme de nature ondulatoire présentent des caractéristiques corpusculaires. C'est ce qu'en somme Planck et Einstein avaient déjà montré, lorsqu'ils avaient démontré que la lumière est constituée de quanta, apparentés en quelque sorte à des particules. En 1923, Compton, ce même Compton qui devait plus tard établir la nature corpusculaire du rayonnement cosmique (voir chapitre 7), démontra que les quanta de lumière possèdent effectivement les mêmes caractéristiques que les particules dont on avait jusque-là l'habitude. Compton montra en effet que des rayons X, lorsqu'ils sont diffractés par la matière, subissent une diminution d'énergie (qui se traduit par une augmentation de leur longueur d'onde). Or, c'est précisément ce que l'on est en droit d'attendre d'une particule ordinaire venant heurter une autre particule ; la vraie particule est poussée en avant et acquiert de l'énergie ; la « fausse » particule, elle, fait demi-tour et perd de l'énergie. L'effet Compton, comme on l'appelle, a joué un grand rôle dans l'histoire de la dualité onde-corpuscule.

L'existence d'ondes de matière est extrêmement importante du point de vue théorique, car c'est de cette façon que l'on a finalement réussi à résoudre le problème de la structure de l'atome.

En 1913, Niels Bohr avait donné de l'atome d'hydrogène une description faisant appel à la toute récente théorie des quanta. Selon cette description, l'atome était constitué d'un noyau autour duquel un électron décrivait l'une ou l'autre d'un certain nombre d'orbites. Ces orbites correspondaient à des figures circulaires bien déterminées dans l'espace ; en « tombant » d'une orbite à l'autre, l'électron perdait de l'énergie qui était émise sous forme de quantum de lumière, caractérisé par une longueur d'onde bien déterminée. A l'inverse, un électron qui effectuait un saut depuis une orbite interne vers une orbite plus périphérique absorbait pour ce faire de l'énergie, sous forme de quantum de lumière dont la longueur d'onde correspondait juste à l'énergie requise pour effectuer le saut. L'atome d'hydrogène ne pouvait donc absorber ou émettre que des quantités d'énergie bien déterminées, celles correspondant précisément aux fréquences des raies caractéristiques de son spectre. Le modèle de Bohr, perfectionné à plusieurs reprises durant les dix années suivantes tout particulièrement par le physicien allemand Arnold Johannes Wilhelm Sommerfeld, qui remplaça les orbites que Bohr avait supposées circulaires par des orbites elliptiques, permettait d'expliquer de façon satisfaisante un grand nombre des caractéristiques observées sur les spectres des divers éléments. Cette théorie valut à Bohr le prix Nobel de physique 1922, et les physiciens allemands James Franck et Gustav Ludwig Hertz (neveu de Heinrich Hertz), dont les travaux expérimentaux avaient permis à Bohr de développer sa théorie, reçurent le prix Nobel de physique en 1925.

Mais la théorie de Bohr ne permettait pas de comprendre pourquoi les orbites occupaient dans l'espace la position qu'elles occupaient ; ces orbites avaient été simplement sélectionnées par Bohr de manière à ce que les prévisions concernant l'absorption et l'émission de lumière soient en accord avec les résultats expérimentaux.

En 1926, le physicien allemand Erwin Schrödinger proposa une autre description de l'atome fondée sur les considérations de de Broglie relatives à la nature ondulatoire des particules de matière. Assimilant donc l'électron à une onde, Schrödinger proposa de traiter l'électron non pas comme une planète gravitant autour du Soleil, mais comme une onde recouvrant le noyau, et présente en quelque sorte en tous les points de l'orbite simultanément. Il s'avéra que si l'on prenait pour longueur d'onde de l'électron celle que lui assignait de Broglie, les orbites de Bohr correspondaient à un nombre entier de ces longueurs d'onde. Dans l'espace entre deux orbites, cette condition n'était pas remplie et les ondes s'additionnaient sans relation de phase ; si bien que seules les orbites de Bohr présentaient la propriété de stabilité.

Schrödinger élabora alors une théorie mathématique de l'atome, qu'il baptisa du nom de *mécanique ondulatoire* ou *mécanique quantique*. Cette théorie supplanta vite l'ancienne théorie de Bohr ; elle était en effet beaucoup plus satisfaisante. Schrödinger partagea le prix Nobel en 1933 avec Dirac, l'auteur de la théorie des antiparticules (voir chapitre 7). Dirac avait lui aussi apporté sa contribution à l'élaboration de cette nouvelle vision de l'atome. Le physicien allemand Max Born reçut également le prix Nobel (en 1954) pour sa participation à l'élaboration de la théorie quantique.

LES RELATIONS D'INDÉTERMINATION DE HEISENBERG

Au fur et à mesure que le temps passait, l'électron apparaissait comme quelque chose de plus en plus vague. De fait, les choses ne devaient pas s'arranger de si tôt, car bientôt les travaux du physicien allemand Werner Heisenberg plongèrent le monde de la physique dans un océan d'incertitude.

Heisenberg avait lui-même élaboré un modèle d'atome dans lequel il renonçait à décrire l'atome en termes d'ondes ou de particules. Selon lui, toute tentative en vue de décrire l'atome dans les mêmes termes que ceux qui nous servent à rendre compte du monde qui nous entoure ne pouvait que manquer son but. Seuls les niveaux d'énergie des diverses orbites étaient susceptibles d'être représentés par des nombres, et toute image devait être bannie. Comme son modèle mathématique fait grand usage de ces objets mathématiques que sont les matrices, sa théorie est généralement connue sous le nom de *mécanique des matrices.*

Heisenberg reçut le prix Nobel en 1932 pour ses travaux dans ce domaine ; néanmoins sa théorie ne connut pas auprès des physiciens le même succès que celle de Schrödinger. Cette dernière, en effet, paraissait devoir rendre les mêmes services que la théorie de Heisenberg, tout en faisant l'économie des abstractions propres à la théorie de Heisenberg : même pour un physicien, il n'est guère facile de travailler sans représentation imagée de ce dont il parle.

L'histoire semblait en 1944 donner raison aux physiciens. En effet, le mathématicien américain d'origine hongroise John von Neumann venait de développer une argumentation au terme de laquelle il apparaissait que la mécanique ondulatoire et la mécanique des matrices étaient équivalentes du point de vue mathématique : tout ce que l'une des théories était capable de démontrer pouvait également l'être en utilisant l'autre. Ne devait-on pas dans ces conditions choisir la moins abstraite des deux théories ?

Mais revenons quelque peu en arrière. Après avoir introduit la mécanique des matrices, Heisenberg avait tenté de traiter le problème de la position de l'électron dans l'atome. Comment peut-on déterminer la position d'une particule ? La réponse qui vient immédiatement à l'esprit est évidemment qu'il suffit de l'observer. Bien. Imaginons donc que nous disposions d'un microscope qui permette d'observer la particule. On peut par exemple imaginer un dispositif qui envoie au bon moment un éclair lumineux, ou tout autre type de radiation, sur l'électron et permette de l'observer. Oui, mais un électron est une particule tellement légère que même un photon peut le déplacer en venant le heurter. Autrement dit, au moment même où nous tentons de mesurer la position de l'électron, nous modifions celle-ci.

C'est là un problème qui n'a rien de nouveau et se produit à chaque instant dans la vie de tous les jours. Quand nous mesurons la pression d'un pneu à l'aide d'un manomètre, nous ne pouvons empêcher qu'un peu d'air s'échappe du pneu, ce qui modifie la pression. De même, lorsque nous plaçons un thermomètre dans une baignoire pour mesurer la température de l'eau, le thermomètre absorbe un peu de la chaleur de l'eau, ce qui modifie la température. De même encore, un ampèremètre dévie une partie du courant qu'il est censé mesurer. On pourrait multiplier les exemples.

Mais, dans tous les cas de mesures ordinaires, la modification apportée par l'appareil de mesure est tellement faible que nous pouvons à bon droit la négliger. Il n'en va pas de même lorqu'il s'agit d'un électron ; car dans ce cas, l'appareil de mesure est au moins aussi grand que la chose à mesurer (on ne peut pas utiliser un appareil qui soit plus petit qu'un électron !). D'où il découle que l'acte de mesure a, dans ce domaine, un effet qui, loin d'être négligeable, est au contraire décisif ! On pourrait évidemment songer à immobiliser l'électron pour pouvoir effectuer la mesure de sa position à un instant donné ; mais alors, nous serions obligés de renoncer à connaître la vitesse qu'il avait au moment où nous l'avons arrêté. On pourrait, à l'inverse, imaginer d'enregistrer les variations de la vitesse de l'électron ; mais alors nous ne pourrions pas déterminer la position de l'électron à chaque instant.

Heisenberg avait donc montré la chose suivante : il n'est pas possible de concevoir un appareil destiné à mesurer la position d'une particule subatomique qui autorise en même temps une détermination précise de l'état de mouvement de la particule. A l'inverse, on ne peut pas déterminer de façon précise le mouvement de la particule sans renoncer de ce fait à une détermination exacte de sa position. En un mot, on ne peut pas déterminer à la fois la position et la vitesse d'une particule à un instant donné.

Mais si Heisenberg avait raison, alors il devait exister, même au zéro absolu, une énergie résiduelle. En effet, imaginons que l'énergie s'annule au zéro absolu et que les particules s'immobilisent complètement, alors seule la position resterait à déterminer (puisque la vitesse aurait la valeur zéro), en contradiction avec le principe énoncé plus haut. C'est donc qu'à 0 °, il reste encore une certaine énergie (appelée *énergie de point zéro*), laquelle maintient un certain degré d'agitation parmi les particules, et donc assure en quelque sorte leur caractère incertain et vague. C'est cette énergie de point zéro, énergie que l'on ne peut pas ôter au système quoi que l'on fasse, qui est responsable de la fluidité de l'hélium au zéro absolu (voir le chapitre 6).

En 1930, Einstein montra que le principe d'indétermination (selon lequel il est impossible de réduire la marge d'erreur sur la position sans du même coup augmenter l'erreur sur la vitesse), implique qu'il est également impossible de diminuer la barre d'erreur sur l'énergie sans du même coup augmenter l'imprécision sur la détermination de l'intervalle de temps dont on dispose pour effectuer la mesure. Einstein avait même tenté d'utiliser cette conclusion pour réfuter le principe de Heisenberg ; mais Bohr avait réussi à montrer que l'argumentation d'Einstein ne tenait pas.

Pourtant, la version du principe de Heisenberg introduite par Einstein s'est avérée d'une utilité remarquable. Elle implique en effet qu'il est possible, lors d'un processus subatomique, que la loi de conservation de l'énergie soit violée pendant un instant très bref – à condition toutefois que tout revienne en l'état initial au bout du compte. Plus l'écart par rapport à la loi de conservation de l'énergie est grand, plus le temps pendant lequel cette loi n'est pas vérifiée doit être court (c'est cette idée qui est à la base de la théorie de Yukawa, voir chapitre 7). La version d'Einstein du principe d'indétermination permet même d'interpréter certains phénomènes du domaine subatomique ; on suppose que certaines particules sont créées à partir de rien et n'existent qu'un temps très court, si bien qu'elles disparaissent avant même qu'on ait pu les détecter. Ces

particules sont dites *virtuelles*, et la théorie des particules virtuelles est l'œuvre de trois hommes : les physiciens américains Julian Seymour Schwinger et Richard Phillips Feynman, et le physicien japonais Shin' Ichirô Tomonaga ; tous trois ont reçu le prix Nobel en 1965.

Depuis 1976 on pense que l'Univers a d'abord été une particule virtuelle, de très petite taille et de très grande masse, que cette particule s'est enflée très rapidement, puis est passée au stade de l'existence. Selon cette conception, l'Univers s'est donc formé à partir de rien, ce qui autorise à spéculer sur la possibilité pour que d'autres univers se forment également à partir de rien.

Le principe d'indétermination a profondément marqué le cours des réflexions des physiciens et des philosophes, tout particulièrement en ce qui concerne le problème de la *causalité* (c'est-à-dire le lien qui unit l'effet à la cause). Mais ce principe n'a pas eu sur le cours de la physique l'effet qu'on lui prête généralement. On lit quelquefois que le principe de Heisenberg prouve le caractère incertain de la nature et montre que la science, après tout, ne peut prétendre répondre à tout, que la connaissance scientifique est à la merci d'un éventuel caprice de l'Univers qui déciderait un beau jour que l'effet ne suit pas forcément la cause. Quelle que soit l'opinion que les philosophes peuvent avoir sur la question, ce qui est sûr, c'est que le principe d'indétermination n'a en rien modifié l'attitude des chercheurs à l'égard de la recherche scientifique. S'il est vrai, par exemple, qu'il est impossible de prédire avec certitude le comportement de chacune des molécules d'un gaz, il n'en reste pas moins que ces molécules, en moyenne, obéissent à des lois déterminées et que leur comportement peut parfaitement être prévu sur la base d'arguments statistiques, de la même façon que les compagnies d'assurance utilisent des tableaux statistiques de mortalité parfaitement fiables, en dépit du fait que la date de la mort d'un individu donné est imprévisible.

De fait, dans la plupart des cas, l'indétermination est d'un ordre de grandeur radicalement différent de celui de la grandeur à mesurer, et peut donc être légitimement négligée, du moins dans les applications pratiques. Rien ne s'oppose à ce que la position et la vitesse d'une étoile, d'une planète, d'une balle de tennis ou d'un grain de sable soient simultanément mesurées, et ce avec toute la précision souhaitable.

Quant à l'indétermination qui intervient dans le domaine quantique, loin d'entraver la tâche du physicien, elle la facilite au contraire. Le principe d'indétermination a permis d'expliquer certaines caractéristiques des phénomènes de radioactivité ou d'absorption de rayonnement par les atomes qui, sans cela, n'auraient pu être interprétés.

Le principe d'indétermination prouve simplement que l'Univers est beaucoup plus complexe que nous l'avions imaginé, mais certainement pas qu'il est irrationnel.

Chapitre 9

La machine

Le feu et la vapeur

Jusqu'ici, je ne me suis intéressé dans ce livre qu'à la science *pure*, c'est-à-dire celle qui a pour but l'explication de l'Univers qui nous entoure. Depuis la nuit des temps, cependant, les hommes ont mis à profit les phénomènes naturels de l'Univers pour améliorer leur sécurité, leur confort et leur plaisir. Ils ont utilisé ces phénomènes, au début, sans y rien comprendre, mais ils en sont venus à les maîtriser grâce à de soigneuses observations, en utilisant leur sens commun, mais aussi en tâtonnant, par essais et erreurs. Une telle application des phénomènes naturels à la satisfaction des besoins humains relève de la technique et a précédé la science.

Mais la science, en se développant, permet à la technique de progresser de plus en plus vite. Dans les temps modernes, la science et la technique sont devenues tellement imbriquées (la science faisant avancer la technique en élucidant les lois de la nature, et la technique faisant avancer la science en fournissant aux scientifiques de nouveaux instruments et dispositifs) qu'il n'est plus possible de vraiment les séparer.

LES DÉBUTS DE LA TECHNIQUE

Si nous retournons aux origines, il nous faut remarquer que, bien que la première loi de la thermodynamique exclue la création d'énergie *ex nihilo*, aucun principe n'empêche la conversion d'une forme d'énergie en

une autre. Toute notre civilisation s'est construite sur la découverte de nouvelles sources d'énergie et leur maîtrise par des moyens toujours plus efficaces et sophistiqués. En fait, la plus grande découverte de l'histoire des hommes a été celle des méthodes permettant de convertir l'énergie chimique d'un carburant tel que le bois en chaleur et en lumière.

C'est il y a environ un demi-million d'années que nos ancêtres hominiens « découvrirent » le feu, bien avant l'apparition de *Homo sapiens* (notre espèce actuelle). Nul doute qu'ils aient rencontré auparavant des feux de broussailles allumés par la foudre et les incendies de forêts, et qu'ils se soient enfuis à cette vue. La découverte des propriétés du feu a dû survenir lorsque la curiosité l'emporta sur la peur.

Un jour, un primitif – probablement un enfant – a peut-être été attiré par les restes de braises d'un feu accidentel et s'est amusé à jouer avec, en y faisant brûler des bouts de bois et en admirant la danse des flammes. Sans aucun doute, les adultes ont dû arrêter ce jeu dangereux jusqu'à ce que l'un d'eux, plus imaginatif, reconnaisse les avantages qu'il y avait à maîtriser la flamme, faisant d'un amusement d'enfant une technique d'adulte. La flamme offrait de la lumière dans l'obscurité et de la chaleur dans le froid. Le feu éloignait les prédateurs. Et il se peut que l'on ait trouvé que sa chaleur rendait la nourriture plus facilement consommable et améliorait son goût (elle tuait aussi les germes et les parasites, mais les hommes préhistoriques ne le savaient pas).

Pendant des centaines de milliers d'années, les êtres humains n'ont pu utiliser le feu qu'en le maintenant constamment allumé. Si un feu s'éteignait accidentellement, ce devait être l'équivalent d'une coupure électrique générale dans nos sociétés modernes. Il fallait alors emprunter du feu à une autre tribu, ou attendre que la foudre tombe. Ce n'est que plus tard que les humains ont appris à obtenir une flamme à volonté ; le feu était alors réellement dompté (figure 9.1). L'*Homo sapiens* y parvint à l'époque préhistorique mais nous ne savons pas exactement quand, où ni comment, et nous l'ignorerons peut-être toujours.

Dès le début de la civilisation le feu servait non seulement à l'éclairage, au chauffage, à la cuisson des aliments et à la protection contre les prédateurs, mais encore à extraire des métaux de leurs minerais et à les travailler ; il servait aussi à cuire les poteries et les briques, et même à faire du verre.

D'autres découvertes importantes ont jalonné la naissance de la civilisation. Vers 9000 av. J.-C., les hommes commencèrent à domestiquer les plantes et les animaux ; c'était le début de l'agriculture et de l'élevage ; ils augmentaient ainsi la quantité disponible de nourriture et trouvaient, dans les animaux, une source directe d'énergie. Les bœufs, les ânes, les chameaux et parfois les chevaux (sans compter les rennes, les yaks, les buffles d'eau, les lamas et les éléphants dans divers endroits du globe), grâce à leurs muscles puissants, permettaient d'accomplir les tâches nécessaires tout en ne sacrifiant que de la nourriture trop grossière pour être mangée par les hommes.

Vers 3500 av. J.-C., on inventa la roue (peut-être, au début, sous la forme d'une roue de potier). En quelques siècles, certainement dès 3000 av. J.-C., on plaça des roues sous les traîneaux, de façon à faire rouler les charges que l'on avait traînées jusqu'alors. Les roues ne constituaient pas une source directe d'énergie, mais elles permettaient de perdre beaucoup moins d'énergie sous forme de frottement.

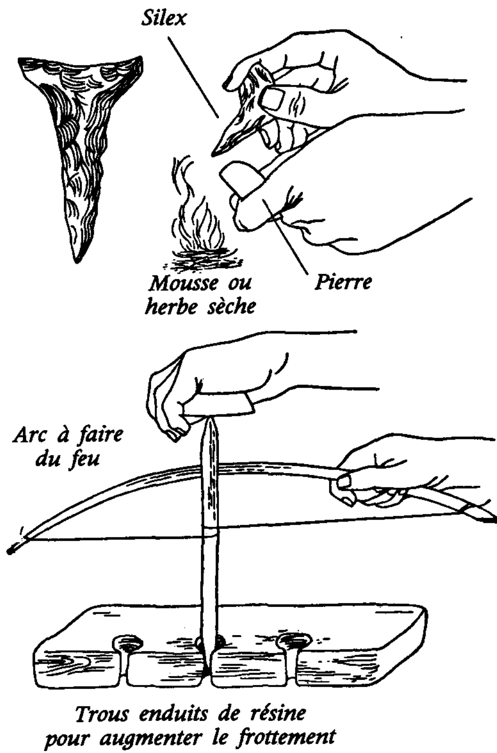


Figure 9.1. Les premières méthodes utilisées pour faire du feu.

C'est aussi vers cette époque qu'on utilisa des radeaux ou des pirogues primitifs pour transporter des charges à l'aide de l'énergie de l'eau courante. Vers 2000 av. J.-C., on se servait des voiles pour capter le vent, afin que l'air en mouvement puisse accélérer le transport ou même forcer le bateau à remonter un faible courant. Vers 1000 av. J.-C., les Phéniciens et leurs voiliers avaient sillonné en tous sens la mer Méditerranée.

Les Romains commencèrent vers 50 av. J.-C. à utiliser des roues hydrauliques. On pouvait, avec un courant d'eau rapide, faire tourner une roue, qui pouvait à son tour en entraîner d'autres pour faire un travail – moudre le grain, écraser le minerai, pomper de l'eau, etc. On commença aussi, à cette époque, à utiliser des moulins à vent : c'est alors la force du vent et non plus celle de l'eau qui faisait tourner la roue. (Les courants d'eau rapides sont rares, tandis qu'il y a du vent presque partout.) Au Moyen Age, les moulins à vent constituaient une importante source d'énergie en Europe de l'Ouest. C'est aussi au Moyen Age que les hommes ont commencé à brûler dans des fours métallurgiques la roche noire qu'on appelle *charbon*, à utiliser l'énergie magnétique dans le compas des bateaux (ce qui a peut-être rendu possibles les grands voyages d'exploration) et à utiliser l'énergie chimique pour les armes.

La première utilisation de l'énergie chimique à des fins destructrices (à l'exception de la pratique sommaire consistant à enflammer les pointes

de flèches) survint vers 670 apr. J.-C., quand un alchimiste syrien du nom de Callinicus inventa le *feu grégeois*, une bombe incendiaire primitive à base de soufre et de naphte, dont on dit qu'elle permit de sauver Constantinople lors de son premier siège par les Musulmans en 673. La poudre à canon fut introduite en Europe au XIII[e] siècle. Roger Bacon la décrivit en 1280, mais elle était connue en Asie depuis des siècles ; elle a dû être introduite en Europe lors des invasions mongoles qui ont débuté vers 1240. Quoi qu'il en soit, c'est au XIV[e] siècle qu'entra en service en Europe l'artillerie actionnée par la poudre et on pense que les canons sont apparus pour la première fois lors de la bataille de Crécy en 1346.

La plus importante des inventions du Moyen Age est celle due à l'Allemand Johannes Gutenberg. Vers 1450, il fondit le premier caractère mobile d'imprimerie, introduisant ainsi dans les affaires humaines cette force puissante qu'est l'imprimerie. Il mit aussi au point l'encre d'imprimerie, où le noir de carbone était en suspension dans l'huile de lin et non, comme auparavant, dans l'eau. Jointes au remplacement du parchemin par le papier (qui, d'après la tradition, avait été inventé par un eunuque chinois, Ts'ai Lun, vers l'an 50, et introduit en Europe par les Arabes au XIII[e] siècle), ces deux inventions ont rendu possibles la production et la distribution à grande échelle de livres et autres œuvres écrites. Aucune invention, avant les temps modernes, ne fut adoptée aussi rapidement. En une génération on imprima environ 40 000 livres.

Dès lors, le trésor de la connaissance humaine n'était plus enfoui dans le secret des collections royales de manuscrits, mais rendu accessible à tous ceux qui pouvaient lire. On commença à écrire des pamphlets, qui constituèrent la première forme d'expression de l'opinion publique. (L'imprimerie est en grande partie la cause du succès de la révolte de Martin Luther contre la papauté en 1517 ; sans elle, ce serait peut-être resté une querelle privée entre moines.)

Ce fut l'imprimerie qui créa l'un des principaux instruments qui ont donné naissance à la science telle que nous la connaissons aujourd'hui : la circulation libre et large des idées. La science était jusqu'alors restée l'objet de communications personnelles réservées à un petit nombre d'initiés ; elle devint alors un champ d'activités immense, impliquant de plus en plus de participants qui formèrent finalement une *communauté scientifique* internationale ; on assista à une mise à l'épreuve rapide et critique des théories nouvelles, ce qui ouvrit sans cesse de nouveaux domaines à la connaissance.

LA MACHINE À VAPEUR

Le grand virage dans la maîtrise de l'énergie se produisit à la fin du XVII[e] siècle. Toutefois, on en avait déjà vu des signes avant-coureurs dans les temps anciens. En effet, pendant l'un des premiers siècles apr. J.-C. (on n'a pu dater sa vie avec plus de précision), l'inventeur grec Héron d'Alexandrie construisit de nombreux dispositifs fonctionnant grâce à l'énergie de la vapeur. Ainsi, Héron utilisait la force de poussée de la vapeur pour ouvrir des portes de temples, faire tourner des sphères, etc. Le monde ancien, alors en pleine décadence, ne put suivre cette avance prématurée.

C'est plus de quinze siècles plus tard que la société, renouvelée et en forte expansion, sut mettre à profit une seconde chance. Celle-ci naquit de la nécessité pressante où l'on était de pomper l'eau des mines, que l'on

creusait toujours plus profondément. La vieille pompe à main (voir le chapitre 5) utilisait le vide pour élever l'eau ; au fur et à mesure que s'écoulait le XVII[e] siècle, les hommes apprécièrent avec de plus en plus d'acuité la grande puissance associée au vide (ou mieux, la puissance de la pression atmosphérique mise en jeu par l'existence du vide).

En 1650, par exemple, le physicien allemand (et maire de la cité de Magdebourg) Otto von Guericke inventa une pompe à air actionnée par la force musculaire de l'homme. Il imagina de coller l'un contre l'autre deux hémisphères de métal munis de flasques, et de pomper l'air compris entre ceux-ci à travers un ajutage fixé sur l'un des hémisphères. Au fur et à mesure que la pression baissait à l'intérieur, la pression atmosphérique extérieure, qui n'était plus contrebalancée intégralement, pressait les hémisphères l'un contre l'autre de plus en plus fort. Finalement, deux attelages de chevaux tirant en sens inverse n'arrivèrent pas à séparer les hémisphères. Cette expérience eut lieu devant une assistance choisie, comprenant même l'empereur d'Allemagne en personne, et elle connut un grand retentissement.

Plusieurs inventeurs se demandèrent alors : pourquoi ne pas utiliser la vapeur au lieu de la force musculaire pour faire le vide ? Supposons que l'on remplisse un récipient d'eau, et qu'on chauffe l'eau jusqu'à ébullition. La vapeur, lors de sa formation, fait sortir l'eau du récipient. Si l'on refroidit alors celui-ci (par exemple en arrosant ses parois d'eau froide), la vapeur restant à l'intérieur du récipient se condense sous la forme de gouttes d'eau, créant ainsi le vide. L'eau que l'on veut pomper (celle d'une mine inondée par exemple) peut alors pénétrer grâce à une valve dans le récipient vidé.

Un physicien français, Denis Papin, vit l'intérêt de la force de la vapeur dès 1679. Il mit au point une *marmite à pression*, dans laquelle de l'eau était portée à ébullition dans un récipient à fermeture hermétique. La vapeur, en s'accumulant, créait une pression qui élevait le point d'ébullition de l'eau et, à cette température plus élevée, la cuisson des aliments se faisait plus vite et mieux. La pression de la vapeur dans la marmite semble avoir donné à Papin l'idée de faire travailler la vapeur. Il plaça un peu d'eau au fond d'un tube et, en chauffant celui-ci, convertit l'eau en vapeur. Celle-ci, en sortant avec force, poussait un piston devant elle.

La première personne à être passée de cette idée à la réalisation pratique d'un dispositif fut un ingénieur militaire anglais du nom de Thomas Savery. Sa *machine à vapeur* pouvait être utilisée pour pomper l'eau d'une mine ou d'un puits ou pour entraîner une roue hydraulique ; on l'appela donc « l'amie du mineur ». Mais elle était dangereuse (la haute pression de la vapeur pouvait faire sauter les tuyaux et les récipients) et d'un mauvais rendement (la chaleur de la vapeur était perdue à chaque refroidissement du cylindre). Savery fit breveter son engin en 1698 et sept ans après, un forgeron anglais, Thomas Newcomen, construisit une machine améliorée qui fonctionnait à basse pression de vapeur ; elle possédait un piston dans un cylindre et employait la pression atmosphérique pour ramener le piston vers le bas.

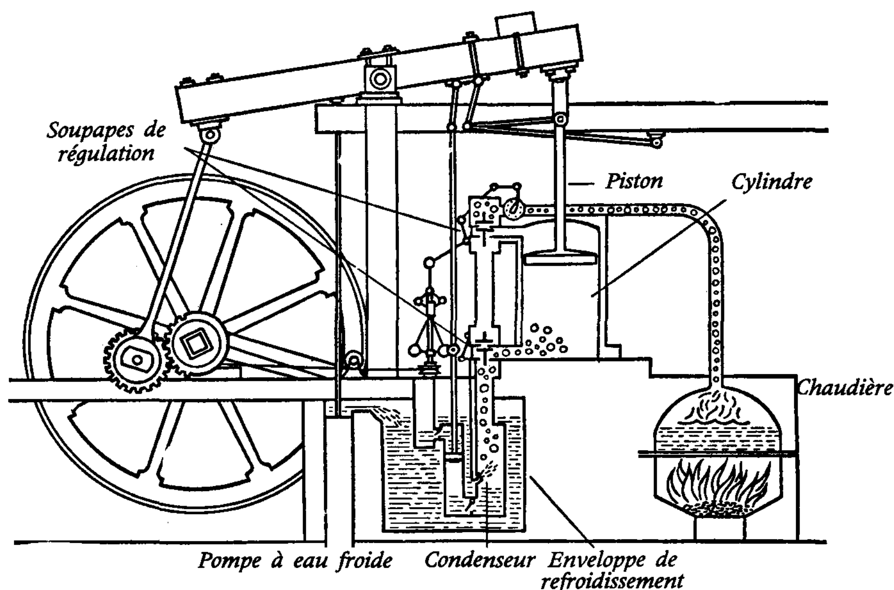
La machine de Newcomen n'était pas, elle non plus, d'un rendement extraordinaire (on refroidissait toujours le cylindre après chaque chauffage), et la machine à vapeur resta une sorte de gadget pendant plus de soixante ans, jusqu'à ce que James Watt, un fabricant écossais d'instruments scientifiques, trouve le moyen d'augmenter son rendement.

Watt avait été engagé par l'université de Glasgow pour réparer un modèle de machine de Newcomen qui ne fonctionnait pas correctement ; il se mit à penser à la consommation abusive de carburant qu'entraînait le dispositif. Pourquoi, après tout, fallait-il refroidir à chaque fois le récipient à vapeur ? Pourquoi ne pas maintenir la température de la chambre à vapeur pendant toute l'opération et envoyer la vapeur dans une deuxième enceinte, une enceinte de condensation que l'on maintiendrait froide ? Watt s'empressa d'ajouter nombre d'autres perfectionnements : emploi de la pression de la vapeur pour aider à pousser le piston, conception d'un ensemble de liaisons mécaniques pour maintenir en ligne droite le mouvement du piston, liaison du va-et-vient du piston à une manivelle qui entraîne une roue, etc. Vers 1782 sa machine à vapeur qui, à partir de la même quantité de charbon, pouvait fournir au moins trois fois plus de travail que celle de Newcomen, était prête à devenir une source universelle de force motrice (figure 9.2).

Après Watt, on augmenta continuellement le rendement de la machine à vapeur, principalement en utilisant de la vapeur de plus en plus chaude à une pression de plus en plus élevée. La fondation par Carnot de la thermodynamique (voir le chapitre 8) trouva pour beaucoup son origine dans le fait que le rendement maximal de tout moteur thermique est proportionnel à la différence de température entre la source chaude (celle de la vapeur) et la source froide.

Au cours du XVIII[e] siècle, on inventa de nombreux dispositifs pour filer et tisser de manière plus rentable. (Ils remplaçaient le rouet qui datait du Moyen Age.) Au début ces dispositifs étaient entraînés par la traction animale ou une roue hydraulique ; mais vers 1790 survint l'étape cruciale consistant à entraîner tout cela avec la machine à vapeur.

Figure 9.2. La machine à vapeur de Watt.



Ainsi, les nouvelles usines textiles que l'on construisait n'avaient plus besoin d'être situées à proximité de cours d'eau rapides ni d'entretenir des animaux. On pouvait les construire partout. L'Angleterre commença à présenter tous les signes d'une révolution sociale quand les classes laborieuses quittèrent la campagne et abandonnèrent l'artisanat pour s'entasser dans les usines (où les conditions de travail étaient incroyablement cruelles, jusqu'à ce que la société apprenne enfin, comme à regret, que les gens ne devaient pas être traités plus mal que les bêtes).

Le même changement se produisit dans d'autres pays qui avaient adopté la machine à vapeur ; ce fut la Révolution industrielle (une expression introduite en 1837 par l'économiste français Jérôme Adolphe Blanqui).

La machine à vapeur révolutionna totalement les transports. En 1787, l'inventeur américain John Fitch construisit un bateau à vapeur qui fonctionnait, mais ce fut un échec financier, et Fitch mourut inconnu et sans renom. Robert Fulton, promoteur plus habile que Fitch, lança en 1807 son navire à vapeur, le *Clermont,* avec tant de publicité et de support financier qu'il en vint à être considéré comme l'inventeur du bateau à vapeur, bien qu'en fait il ne soit pas plus le constructeur du premier bateau à vapeur que Watt celui de la première machine à vapeur.

Le nom de Fulton devrait cependant rester dans les mémoires pour son opiniâtreté à vouloir construire un vaisseau sous-marin. Ses sous-marins n'étaient pas utilisables, mais ils anticipèrent nombre de perfectionnements modernes. Il en construisit un appelé le *Nautilus,* qui inspira probablement à Jules Verne le nom donné au submersible de son roman d'anticipation, *Vingt mille lieues sous les mers,* paru en 1870. Ce nom fut également choisi pour le premier sous-marin à propulsion nucléaire (voir le chapitre 10).

Vers 1830, des bateaux à vapeur traversaient l'Atlantique et étaient propulsés par une *hélice,* considérable progrès par rapport aux roues à aubes du début. Et, vers 1850, les beaux et rapides voiliers Yankee Clippers avaient commencé à amener leurs voiles pour être remplacés par les vapeurs dans les flottes marchandes et militaires du monde entier.

Un peu plus tard, un ingénieur britannique, Charles Algernon Parsons (un fils de Lord Rosse, qui avait découvert la Nébuleuse du Crabe), pensa à un perfectionnement capital de la machine à vapeur pour la propulsion des navires. Au lieu de faire entraîner par la vapeur un piston qui, à son tour, entraînait une roue, il pensa à éliminer cet intermédiaire : il suffisait de diriger un jet de vapeur sur des ailettes placées à la périphérie d'une roue pour la faire tourner. La roue devait supporter de hautes températures et de grandes vitesses ; en 1884, Parsons réalisa la première *turbine à vapeur* opérationnelle.

En 1897, au jubilé de diamant de la reine Victoria, la marine britannique présentait une imposante revue de ses navires de guerre à vapeur quand le bateau à turbine à vapeur de Parsons, le *Turbinia,* les dépassa en silence à une vitesse de 35 nœuds. Aucun bateau de la marine britannique n'aurait pu le rattraper ; ce fut la meilleure publicité qu'on ait pu imaginer. Il ne fallut pas longtemps pour que les vaisseaux de guerre et de commerce soient propulsés par turbine à vapeur.

Pendant ce temps, la machine à vapeur avait aussi commencé à dominer le transport terrestre. En 1814, l'inventeur anglais George Stephenson construisit la première *locomotive à vapeur* mise en service (grâce en grande partie au travail antérieur d'un autre ingénieur anglais, Richard

Trevithick). Le mouvement de va-et-vient de pistons entraînés par la vapeur pouvait faire tourner des roues métalliques sur des rails d'acier, comme il pouvait faire tourner les roues à aubes sur l'eau. En 1830, l'industriel américain Peter Cooper construisit la première locomotive du continent. Pour la première fois dans l'histoire, le transport terrestre devenait aussi commode que celui sur mer, et le commerce par voie de terre pouvait entrer en compétition avec le commerce maritime. Vers 1840, le chemin de fer avait atteint le Mississipi ; vers 1869, tous les États-Unis étaient sillonnés par le rail.

L'électricité

Par nature, la machine à vapeur n'est utilisable que pour une production d'énergie à grande échelle et continue. Elle ne peut fournir d'énergie en petite quantité ou de façon intermittente sur la simple pression d'un bouton ; une « petite » machine à vapeur, dans laquelle on éteindrait ou allumerait le feu à la demande, serait une absurdité. Mais la génération qui a vu le développement de la machine à vapeur a vu aussi la découverte d'un moyen de transformer l'énergie pour en faire justement cela : une source d'énergie qu'on peut amener n'importe où, en petite ou en grande quantité, sur la simple pression d'un bouton. Il s'agit, bien sûr, de l'électricité.

L'ÉLECTRICITÉ STATIQUE OU ÉLECTROSTATIQUE

Le philosophe grec Thalès, vers 600 av. J.-C., notait en ses écrits qu'une résine fossile des rives de la Baltique, que nous nommons ambre et que les Grecs appelaient *elecktron*, avait la propriété d'attirer les plumes, les fils ou autres menus objets quand on la frottait avec un morceau de fourrure. Ce fut l'Anglais William Gilbert, connu par ses recherches sur le magnétisme (voir le chapitre 5), qui suggéra le premier que cette force d'attraction soit nommée *électricité*, à partir du mot grec *elektron*. Gilbert découvrit qu'outre l'ambre, certains autres matériaux comme le verre pouvaient acquérir la propriété électrique quand on les frottait.

En 1733, le chimiste français Charles François de Cisternay Du Fay découvrit que si deux barreaux d'ambre, ou deux barreaux de verre, étaient électrisés par frottement, ils se repoussaient l'un l'autre. Cependant, un barreau de verre électrisé attirait un barreau d'ambre électrisé. Si on les faisait se toucher, ils perdaient tous deux leur électricité. Il fut ainsi amené à penser qu'il y avait deux sortes d'électricité, une *vitreuse* et une *résineuse*.

L'Américain Benjamin Franklin, qui manifestait un immense intérêt pour l'électricité, suggéra qu'il s'agissait en fait d'un seul et même fluide. Quand on frottait le verre, l'électricité y pénétrait, le rendant « chargé positivement » ; au contraire, quand on frottait l'ambre, l'électricité en sortait et l'ambre devenait donc « chargé négativement ». Et quand un barreau négatif entrait en contact avec un positif, le fluide électrique circulait du positif au négatif jusqu'à réaliser un équilibre neutre.

C'était une hypothèse remarquablement hardie. Si nous remplaçons le mot *fluide* de Franklin par le mot *électrons* et si nous renversons la direction du fluide (en réalité les électrons circulent de l'ambre vers le verre), nous constatons que son explication était correcte.

En 1740, un inventeur français, Jean Théophile Désaguliers, suggéra qu'on nomme *conducteurs* les corps à travers lesquels le fluide électrique circule librement (par exemple les métaux) et *isolants* ceux à travers lesquels il ne se déplace pas librement (par exemple le verre et l'ambre).

Les expérimentateurs découvrirent qu'on pouvait accumuler progressivement une forte charge électrique sur un conducteur si celui-ci était isolé par du verre ou une couche d'air. Le dispositif de ce genre le plus spectaculaire fut la *bouteille de Leyde*. C'est l'Allemand Ewald Jürgen von Kleist qui la réalisa le premier en 1745, mais le premier usage pratique en fut fait en Hollande, à l'université de Leyde, où le Hollandais Petrus Van Musschenbroek la construisit indépendamment quelques mois plus tard. La bouteille de Leyde constitue un exemple de ce qu'on appelle aujourd'hui un *condensateur*, c'est-à-dire deux surfaces conductrices, séparées par une fine couche d'isolant, sur lesquelles on peut emmagasiner une grande quantité de charge électrique.

Dans le cas de la bouteille de Leyde, la charge est accumulée sur une feuille d'étain recouvrant une bouteille de verre, au moyen d'une chaîne de laiton qui pénètre dans la bouteille à travers un bouchon. Quand on touche la bouteille chargée, on ressent une forte secousse électrique. La bouteille de Leyde peut aussi produire une étincelle. Naturellement, plus la charge d'un corps est grande, plus grande est sa tendance à s'échapper. La force entraînant les électrons de la région où ils sont le plus nombreux (le *pôle négatif*) vers la région où il y en a le moins (le *pôle positif*) est la *force électromotrice* (fem) ou *potentiel électrique*. Si le potentiel électrique devient assez élevé, les électrons peuvent même sauter un espace isolant entre le pôle négatif et le pôle positif. Ils peuvent ainsi franchir un espace d'air en produisant une étincelle brillante et un bruit de pétard. La lumière de l'étincelle est due au rayonnement émis par les collisions entre les électrons et les molécules d'air ; le bruit résulte de la dilatation de l'air rapidement chauffé, suivie du claquement de l'air plus froid qui se précipite dans les régions de vide partiel momentanément produites.

On en vint tout naturellement à se demander si l'éclair et le tonnerre n'étaient pas le même phénomène, à une plus vaste échelle, que la petite étincelle tirée de la bouteille de Leyde. Un Anglais, William Wall, avait justement émis cette hypothèse en 1708. Cette idée suffit à inspirer à Benjamin Franklin sa fameuse expérience de 1752. Le cerf-volant qu'il fit voler dans un orage était muni d'une pointe métallique à laquelle il avait attaché un fil de soie pouvant conduire au sol l'électricité des nuages d'orage. Quand Franklin mit la main près d'une clé métallique attachée au fil de soie, il y eut une étincelle entre la clé et son doigt (figure 9.3). Franklin fit se recharger la clé à partir des nuages, puis l'utilisa pour charger une bouteille de Leyde ; cette dernière se chargea aussi bien ainsi que par les méthodes habituelles. Franklin apportait par là la preuve que les nuages d'orage étaient chargés d'électricité et que le tonnerre et l'éclair n'étaient au fond qu'un effet de gigantesque bouteille de Leyde dans le ciel, où les nuages formaient un pôle et la Terre l'autre.

Ce qui est extraordinaire c'est que Franklin ait survécu à cette expérience. D'autres, en effet, la tentèrent et trouvèrent la mort, parce que

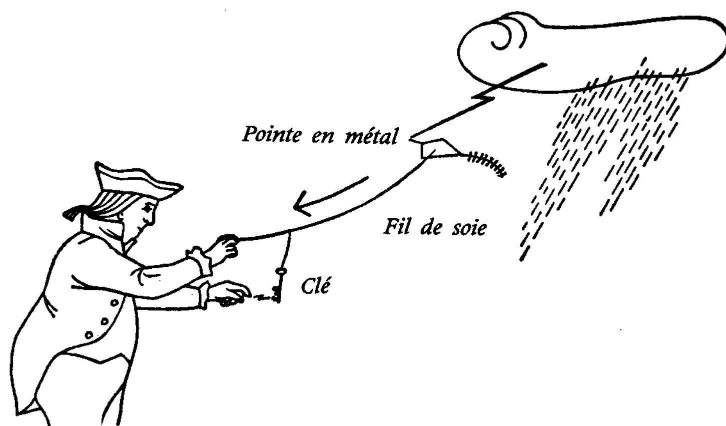


Figure 9.3. L'expérience de Franklin.

la charge induite sur la pointe métallique du cerf-volant s'était accumulée au point de produire une décharge intense et mortelle à travers le corps de l'homme tenant le fil du cerf-volant.

Franklin concrétisa cette avance théorique par une application pratique. Il mit au point le paratonnerre, constitué simplement d'un barreau de fer attaché au plus haut point d'un bâtiment et relié par des fils conducteurs au sol. La pointe aiguisée du barreau collecte les charges électriques des nuages au-dessus d'elle et, si par hasard l'éclair frappe la maison, la charge est transportée sans risque jusqu'au sol.

Les dégâts causés par la foudre diminuèrent de manière remarquable au fur et à mesure qu'on dressait des paratonnerres sur les édifices, tant en Europe que dans les colonies américaines – ce n'était pas un mince résultat. Pourtant, encore aujourd'hui, alors que deux milliards d'éclairs atteignent la planète annuellement, on estime qu'ils tuent vingt personnes et en blessent quatre-vingts par jour.

L'expérience de Franklin eut deux effets électrisants (que l'on me pardonne cet effet facile). D'une part, le monde entier commença tout à coup à s'intéresser à l'électricité. D'autre part, cela inscrivit les colonies américaines sur la carte du monde, culturellement parlant. Pour la première fois, un Américain avait manifesté une aptitude scientifique capable d'impressionner les Européens cultivés du Siècle des Lumières. Quand, vingt-cinq ans plus tard, Franklin représenta à Versailles les États-Unis encore à leur naissance, il conquit aussitôt le respect, non seulement en tant que simple envoyé d'une nouvelle république, mais aussi en tant que génie qui avait dompté la foudre et l'avait conduite humblement à terre. Ce cerf-volant a apporté une contribution non négligeable à la cause de l'indépendance américaine.

A la suite des travaux de Franklin, la recherche dans le domaine de l'électricité avança par bonds successifs. En 1785, le physicien français Charles Augustin de Coulomb réussit à mesurer l'attraction et la répulsion électriques. Il montra que cette attraction (ou répulsion) entre des charges

données varie comme l'inverse du carré de la distance entre charges. De ce point de vue, l'attraction électrique ressemble à l'attraction gravitationnelle. En l'honneur de cette découverte, l'unité de quantité d'électricité du système international se nomme le *coulomb.*

L'ÉLECTRODYNAMIQUE

Peu après, l'étude de l'électricité connut un nouveau tournant aussi soudain que fécond. Je n'ai parlé jusqu'ici que d'électricité statique ou *électrostatique ;* elle concerne des charges électriques placées sur des objets et y restant. La découverte du déplacement des charges électriques, c'est-à-dire des courants électriques, de *l'électrodynamique,* débuta avec l'anatomiste italien Luigi Galvani. En 1791, il découvrit accidentellement que des muscles de cuisses de grenouilles disséquées se contractaient quand on les touchait simultanément avec deux métaux différents (d'où d'ailleurs l'origine du verbe *galvaniser*).

Les muscles se comportaient comme s'ils avaient été excités par une décharge électrique venant d'une bouteille de Leyde ; c'est pourquoi Galvani supposa que les muscles contenaient quelque chose qu'il nomma *électricité animale.* D'autres, cependant, se demandaient si l'origine de la charge électrique ne se trouvait pas plutôt dans la jonction des deux métaux que dans le muscle. En 1880, le physicien italien Alessandro Volta étudia les combinaisons de métaux différents, reliés non par du tissu musculaire mais par des solutions variées.

Au début, Volta utilisa des chaînes de métaux différents reliés par des bols à demi pleins d'eau salée. Puis, pour éviter que le liquide ne soit trop facilement renversé, il prépara de petits disques de cuivre et de zinc qu'il empila en les alternant. Il utilisa aussi des disques de buvard imprégnés d'eau salée, de sorte que la *pile de Volta* consistait en argent, buvard, zinc, argent, buvard, zinc, argent, et ainsi de suite. On pouvait tirer du courant électrique continûment d'un tel assemblage.

On appelle pile tout empilement analogue. L'instrument de Volta était la première *pile électrique* (figure 9.4). Il fallut attendre un siècle pour que les scientifiques comprennent comment les réactions chimiques mettent en jeu des transferts d'électrons, pour qu'ils sachent interpréter les courants électriques en termes de déplacements d'électrons. Jusque-là, ils utilisèrent la notion de courant sans en comprendre tous les détails.

Humphry Davy utilisa un courant électrique pour séparer les atomes composant des molécules fortement liées et réussit le premier, en 1807 et 1808, à préparer des métaux comme le sodium, le potassium, le magnésium, le calcium, le strontium et le baryum. Faraday (assistant et protégé de Davy) élucida les règles générales de cette *électrolyse* qui permet de dissocier les constituants des molécules ; son travail, cinquante ans plus tard, fut un guide précieux pour Arrhenius dans l'élaboration de l'hypothèse de la dissociation (voir le chapitre 5).

Les usages multiples de l'électrodynamique, durant les cent cinquante ans qui suivirent la découverte de la pile Volta, semblaient avoir réduit l'électrostatique à l'état de simple curiosité historique. C'était compter sans le progrès des connaissances et l'ingéniosité des techniciens. Vers 1960, l'inventeur américain Chester Carlson imagina un dispositif pratique pour photocopier des documents en attirant du noir de carbone sur un papier

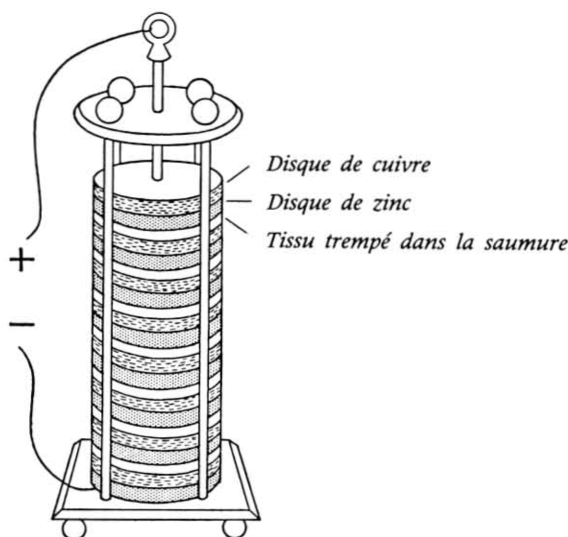


Figure 9.4. La pile de Volta. Les deux métaux différents en contact donnent naissance à un flux d'électrons, qui sont conduits d'un *élément* au suivant par le tissu imprégné de sel. La *pile sèche*, si familière aujourd'hui, à base de carbone et de zinc, a été conçue à l'origine par Bunsen (célèbre spectroscopiste) en 1841.

grâce à des forces électrostatiques bien localisées. On appelle ce système de photocopie sans solutions liquides ni milieux humides *xérographie* (du grec signifiant « écriture sèche ») ; il a révolutionné la vie dans les bureaux.

On a immortalisé les noms des premiers chercheurs dans le domaine de l'électricité en les donnant aux différentes unités servant à mesurer l'électricité. J'ai déjà mentionné le *coulomb* comme unité de charge électrique. Le *faraday* constitue une autre unité : 96 500 coulombs valent 1 faraday. Le nom de Faraday s'est trouvé être deux fois utilisé : le *farad* est une unité de capacité. L'unité d'intensité du courant électrique (la quantité de charge électrique passant à travers un circuit en un temps donné) s'appelle *l'ampère*, d'après le physicien français Ampère (voir le chapitre 5). Un ampère vaut un coulomb par seconde. L'unité de force électromotrice (la force qui entraîne le courant) est le *volt*, d'après Volta.

Une fem donnée n'a pas le même effet d'entraînement de la charge électrique par unité de temps selon qu'elle est appliquée à différents circuits. Elle crée un courant fort à travers de bons conducteurs, un courant faible à travers de mauvais conducteurs, et pratiquement pas de courant si on l'applique à un isolant. En 1827, le mathématicien allemand Georg Simon Ohm étudia cette *résistance* au courant électrique et montra qu'on peut la relier au courant parcourant un circuit sous l'effet d'une fem donnée. La résistance est alors le quotient du nombre de volts par le nombre d'ampères. Ceci constitue la *loi d'Ohm*, et l'unité de résistance électrique est *l'ohm*, 1 ohm valant 1 volt divisé par 1 ampère.

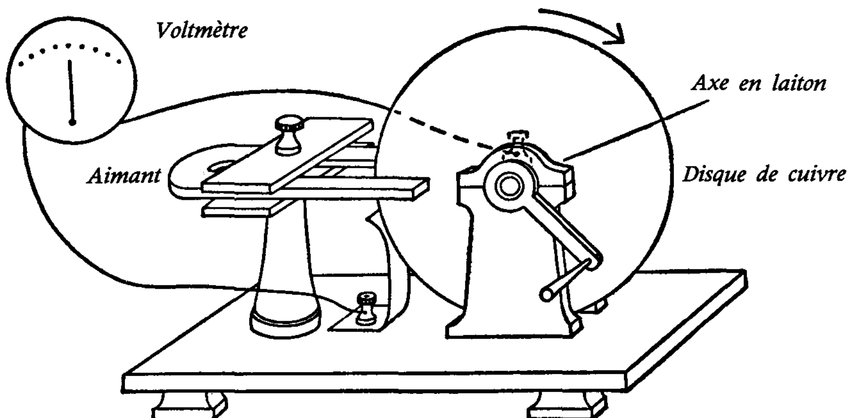
LES GÉNÉRATEURS ÉLECTRIQUES

La conversion d'énergie chimique en énergie électrique, comme dans la pile de Volta et les modèles extrêmement variés de piles qui lui succédèrent, a toujours été relativement onéreuse car les produits chimiques utilisés sont rares ou coûteux. C'est pour cette raison que, bien que l'électricité fût utilisée en laboratoire avec avantage dès le début du XIX[e] siècle, elle ne pouvait l'être dans les applications industrielles à grande échelle.

Il y a eu quelques essais sporadiques pour utiliser comme source d'électricité les réactions chimiques mises en jeu dans la combustion des carburants ordinaires. Des carburants comme l'hydrogène (ou mieux, le charbon) sont beaucoup moins chers que des métaux comme le cuivre et le zinc. Dès 1839, le scientifique anglais William Grove imagina une pile électrique fonctionnant à partir de la réaction hydrogène-oxygène. C'était intéressant mais inutilisable. Ces dernières années, les physiciens ont beaucoup travaillé sur des versions plus utilisables de telles *piles* à combustible. La théorie est complète ; seuls les problèmes pratiques restent à résoudre, et il s'avère qu'ils sont des plus difficiles.

Dans ces conditions, il ne faut pas s'étonner de ce que l'utilisation à grande échelle de l'électricité (dans la deuxième partie du XIX[e] siècle) n'ait pas été réalisée au moyen de piles électriques. Dès 1930, Faraday avait produit de l'électricité en utilisant le mouvement mécanique d'un conducteur à travers les lignes de force d'un aimant (figure 9.5 ; voir aussi le chapitre 5). Dans un tel *générateur électrique*, ou *dynamo* (du mot grec pour « puissance »), l'énergie cinétique du mouvement est convertie en électricité. Ce mouvement est entretenu par une machine à vapeur, chauffée en brûlant un combustible. On peut donc, de manière beaucoup plus indirecte que dans une pile à combustible, convertir en électricité l'énergie de la combustion du charbon (ou même du bois). Vers 1844, des modèles énormes et encombrants de tels générateurs étaient en service et fournissaient l'énergie nécessaire à des machines-outils.

Figure 9.5. La dynamo de Faraday. Le disque tournant de cuivre coupe les lignes de force de l'aimant, induisant ainsi un courant dans le voltmètre.



Ce dont on avait besoin, c'était d'aimants encore plus forts, de façon que le mouvement, à travers des lignes de champ plus intenses, fournisse de plus grands courants éléctriques. Ces aimants plus forts furent obtenus, à leur tour, à l'aide de courants électriques. En 1823, l'expérimentateur anglais William Sturgeon enroula dix-huit tours de fil de cuivre nu autour d'un barreau de fer en forme d'U, réalisant ainsi un électro-aimant. Quand le courant circulait dans le fil, le champ magnétique qu'il produisait était concentré dans le barreau, qui pouvait alors soulever vingt fois son propre poids de fer. Quand le courant était coupé, le barreau n'était plus aimanté et ne soulevait plus rien.

En 1829, le physicien américain Joseph Henry améliora ce dispositif en utilisant du fil isolé. On pouvait enrouler celui-ci en boucles serrées sans craindre les courts-circuits. Chaque boucle augmentait l'intensité du champ magnétique et la puissance de l'électro-aimant. Vers 1831, Henry fabriqua un électro-aimant de taille moyenne capable de soulever une tonne de fer.

L'électro-aimant apparut très vite comme la voie vers de meilleurs générateurs électriques. En 1845, le physicien anglais Charles Wheatstone l'utilisa dans ce but. L'interprétation mathématique que donna Maxwell, vers 1860, des travaux de Faraday (voir le chapitre 5) fournit une meilleure compréhension de la théorie des lignes de champ ; en 1872, l'ingénieur électricien allemand Friedrich von Hefner-Alteneck conçut le premier générateur réellement efficace. L'électricité pouvait enfin être produite à bas prix et en masse, et non seulement à partir de la combustion d'un carburant mais aussi à partir des chutes d'eau.

PREMIÈRES APPLICATIONS TECHNIQUES DE L'ÉLECTRICITÉ

Parmi ceux qui réalisèrent les premières applications techniques de l'électricité, Joseph Henry mérite une mention particulière. Henry commença par inventer le *télégraphe.* Il conçut un système de relais qui permettait de transmettre un courant électrique à travers plusieurs kilomètres de fils. L'intensité du courant (à tension constante) décroît rapidement le long d'un fil très long ; dans les relais inventés par Henry, on utilisait ce signal très fortement atténué pour commander un petit électro-aimant qui lui-même commandait un contact ; on réussissait ainsi à injecter de la puissance dans le circuit, de loin en loin, à des intervalles appropriés. On pouvait alors envoyer un message codé sous la forme d'impulsions électriques à des distances considérables.

Henry était un homme qui se tenait hors du monde. Il pensait que le savoir devait être partagé entre tous, aussi ne faisait-il pas breveter ses inventions ; c'est ainsi qu'il perdit le bénéfice moral de sa découverte. Tout le crédit en alla à un artiste (qui était aussi un bigot excentrique), Samuel Finley Breese Morse. Avec l'aide d'Henry, que celui-ci lui prodiguait généreusement, mais qu'il ne reconnut qu'avec de fortes réticences, Morse construisit le premier télégraphe utilisable en 1844. La principale contribution originale de Morse à la télégraphie fut le système de points et traits constituant l'*alphabet morse.*

La plus importante innovation due à Henry dans le domaine de l'électricité fut le moteur électrique. Il montra que le courant électrique pouvait servir à entraîner une roue, tout comme la rotation d'une roue

pouvait fournir du courant électrique. Et une roue entraînée électriquement (un moteur) pouvait être utilisée pour entraîner des machines. On pouvait installer le moteur n'importe où, le mettre en route ou l'arrêter à volonté (sans avoir à attendre la mise en pression d'une chaudière), le fabriquer aussi petit qu'on le désirait (figure 9.6).

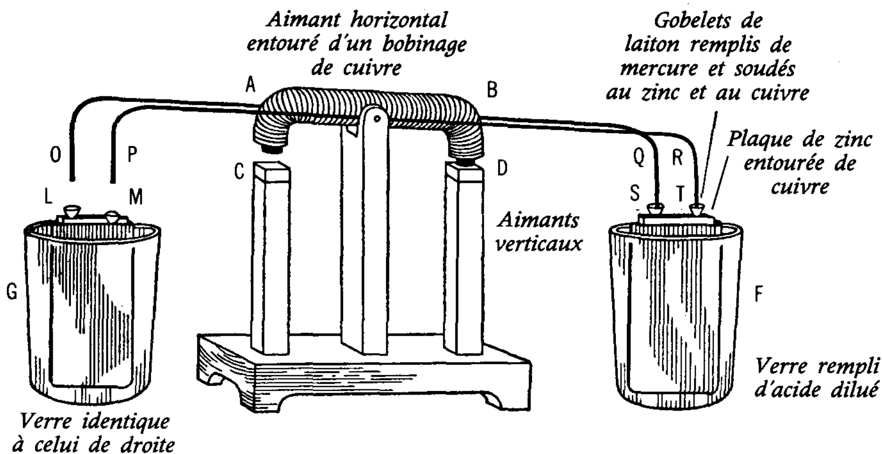
Le problème subsistait de transporter l'électricité de la station génératrice jusqu'à l'endroit où l'on utilisait le moteur. Il fallait trouver un moyen de diminuer la perte d'énergie électrique (sous forme de dissipation de chaleur) lors de son transport à travers les conducteurs.

L'une des solutions fut le *transformateur.* Les expérimentateurs du courant électrique avaient trouvé que l'électricité subit beaucoup moins de pertes si on la transporte avec un débit faible. On portait donc la sortie du générateur à une tension élevée à l'aide d'un transformateur. De cette façon, si on multipliait la tension par trois, par exemple, le transformateur réduisait le courant (le débit) d'un tiers. A la station réceptrice, on diminuait la tension de façon à augmenter le courant pour l'utiliser dans les moteurs.

Le transformateur fonctionne en utilisant le courant *primaire* pour induire un courant à haute tension dans une bobine *secondaire.* Cette induction n'a lieu que si le champ magnétique dans la seconde bobine varie. Comme un courant constant ne peut entraîner cette variation, on utilise un courant variant continuellement ; il augmente dans un sens jusqu'à un maximum, puis diminue jusqu'à s'annuler, et recommence à augmenter dans l'autre sens. On l'appelle *courant alternatif.*

Le courant alternatif ne l'emporta sur le courant continu qu'après une dure bataille. Thomas Alva Edison, le plus illustre des techniciens de l'électricité des dernières décades du XIX[e] siècle, était un tenant du courant

Figure 9.6. Le moteur d'Henry. Le barreau aimanté vertical D attire l'aimant B sur lequel est enroulé du fil, plongeant les longues sondes de métal Q et R dans les gobelets de laiton S et T, qui servent de bornes à la pile humide F. Le courant circule alors dans l'aimant horizontal, produisant un champ magnétique qui rapproche A de C. Le même processus se répète de l'autre côté. La barre horizontale oscille donc.



continu et il créa la première usine électrique à New York en 1882 dans le but de fournir du courant, continu justement, pour alimenter l'éclairage électrique qu'il avait inventé. Il combattit le courant alternatif, arguant qu'il était plus dangereux (il disait, par exemple, que ce courant était utilisé pour la chaise électrique). Il fut contré violemment par Nikola Tesla, un ingénieur qui avait travaillé pour lui et qu'il avait fort mal traité. Tesla mit au point un système à courant alternatif en 1888. En 1893, George Westinghouse, lui aussi tenant du courant alternatif, remporta une victoire cruciale sur Edison. Il obtint pour sa compagnie d'électricité le contrat de développement des usines hydro-électriques des chutes du Niagara. Ces usines fonctionnaient en courant alternatif. Au cours des décades suivantes, Steinmetz établit la théorie des courants alternatifs sur une base mathématique rigoureuse. Aujourd'hui, le courant alternatif est universellement employé dans les systèmes de distribution d'électricité. (En 1966, pour mémoire, des ingénieurs de chez General Electric ont mis au point une sorte de transformateur à courant continu – qu'on avait longtemps tenu pour impossible –, mais il fonctionnait aux basses températures de l'hélium liquide, avec un très faible rendement. Cette innovation était fascinante d'un point de vue théorique, mais guère utilisable d'un point de vue commercial.)

La technique électrique

La machine à vapeur constitue un *moteur primaire :* elle prend l'énergie déjà existante dans la nature (l'énergie chimique du bois, du pétrole ou du charbon) et la convertit en travail. Le moteur électrique n'est pas un moteur primaire : il convertit l'électricité en travail, mais l'électricité doit elle-même provenir de la combustion d'un carburant ou de l'exploitation d'une chute d'eau. C'est pour cette raison que l'électricité est plus onéreuse que la vapeur pour les gros travaux. On peut toutefois l'utiliser dans ce but. A l'exposition de Berlin en 1879, une locomotive propulsée par l'électricité (elle utilisait un troisième rail comme source de courant) tirait un train de voyageurs. Les trains électriques sont très courants aujourd'hui, surtout pour les liaisons rapides entre villes proches, car la dépense supplémentaire est plus que compensée par les avantages de propreté et de souplesse du fonctionnement.

LE TÉLÉPHONE

Mais l'électricité est irremplaçable quand elle permet de réaliser ce que la vapeur ne peut pas. C'est, par exemple, le cas du téléphone, dont le brevet fut pris en 1876 par l'inventeur écossais Alexander Graham Bell. Dans le microphone du téléphone, les ondes sonores émises par celui qui parle heurtent un fin diaphragme d'acier et le font vibrer selon leur propre configuration. Les vibrations du diaphragme, à leur tour, engendrent une configuration analogue dans un courant électrique qui augmente ou décroît exactement comme les ondes sonores. Dans l'écouteur du téléphone, les fluctuations d'intensité du courant agissent sur un électro-aimant qui fait vibrer un diaphragme, ce qui restitue les ondes sonores.

Le premier téléphone était sommaire et marchait à grand-peine ; il fut néanmoins le clou de l'exposition qui se tint en 1876 à Philadelphie pour célébrer le centième anniversaire de la Déclaration d'indépendance. L'empereur du Brésil, Pierre II, qui y était invité, essaya l'appareil et le laissa tomber d'étonnement en s'exclamant : « Ça parle ! », phrase qui fit la une des quotidiens. Un autre visiteur, Kelvin, fut tout aussi impressionné, tandis que le grand Maxwell s'étonna de ce que quelque chose d'aussi simple puisse reproduire la voix humaine. En 1877, la reine Victoria fit l'acquisition d'un téléphone ; le succès était assuré.

Toujours en 1877, Edison apporta à son invention un perfectionnement essentiel. Il construisit un microphone contenant de la poudre de carbone peu tassée. Quand le diaphragme appuyait sur la poudre de carbone, celle-ci conduisait plus de courant ; quand le diaphragme s'éloignait, la poudre conduisait moins de courant. De cette façon, les ondes sonores de la voix étaient converties par le microphone en impulsions électriques avec une grande fidélité et la voix que l'on entendait dans l'écouteur était reproduite avec une clarté très améliorée.

On ne pouvait transmettre les messages téléphoniques bien loin sans des investissements ruineux en fil de cuivre de gros diamètre (donc de faible résistance). A la fin du siècle dernier, le physicien américain d'origine yougoslave Michael Idvorsky Pupin inventa une méthode consistant à charger un fin fil de cuivre par des bobines d'auto-induction placées à intervalles réguliers. Ce procédé renforçait les signaux et permettait de les transporter à longue distance. La compagnie des téléphones Bell acheta le dispositif en 1901 ; vers 1915, la téléphonie à longue distance était un fait bien tangible lors de l'ouverture de la ligne New York-San Francisco.

L'opératrice (c'était presque toujours une femme) du téléphone devint un personnage indispensable et d'une importance croissante durant un demi-siècle, jusqu'à ce que son rôle commence à s'estomper avec les débuts du téléphone à cadran en 1921. L'automatisation continua à se développer, si bien qu'en 1983, aux États-Unis, en dépit d'une grève de deux semaines de centaines de milliers d'employés du téléphone, le service téléphonique a continué de fonctionner sans interruption. Aujourd'hui, faisceaux hertziens et satellites de communication augmentent encore les possibilités du téléphone.

L'ENREGISTREMENT DU SON

En 1877, un an après l'invention du téléphone, Edison fit breveter son *phonographe*. Les sillons des premiers enregistrements étaient gravés sur une feuille d'étain enroulée autour d'un cylindre tournant. L'inventeur américain Charles Sumner Tainter remplaça ce système par des cylindres de cire en 1885, puis en 1887 Émile Berliner introduisit des cylindres recouverts de cire. En 1904, Berliner créa un perfectionnement encore plus important : le disque plat de phonographe sur lequel l'aiguille vibrait latéralement. La plus grande compacité de ce dernier lui permit de remplacer presque aussitôt le cylindre d'Edison (où l'aiguille vibrait de haut en bas).

En 1925, on commença à faire les enregistrements électriquement au moyen d'un *microphone* qui convertissait le son en un courant électrique grâce à un cristal piézo-électrique à la place du diaphragme métallique – le cristal permettant une meilleure qualité de reproduction du son. Dans les années trente, on introduisit l'usage des tubes radio pour l'amplification.

En 1948, le physicien américain d'origine hongroise Peter Goldmark mit au point le *disque microsillon* qui tournait à 33 tours 1/3 par minute au lieu des 78 tours par minute des anciens disques. Un simple disque microsillon pouvait emmagasiner six fois plus de musique que le disque traditionnel et permettait d'écouter des symphonies sans être obligé de retourner et de changer les disques.

L'électronique rendit possible la *haute fidélité* et la *stéréophonie,* ce qui a eu pour effet, en ce qui concerne le son proprement dit, de supprimer tous les obstacles mécaniques entre l'orchestre (ou le chanteur) et l'auditeur.

L'enregistrement magnétique du son fut inventé en 1898 par un ingénieur électricien danois nommé Valdemar Poulsen, mais il dut attendre plusieurs améliorations techniques avant de devenir d'usage courant. Un électro-aimant, excité par un courant électrique traduisant le motif sonore, magnétise une poudre déposée sur un ruban ou un fil se déplaçant devant lui ; la lecture se fait par un électro-aimant qui lit le motif magnétique et le convertit de nouveau en un courant qui reproduit le son.

L'ÉCLAIRAGE ARTIFICIEL AVANT L'ÉLECTRICITÉ

De tous les miracles réalisés par l'électricité, le plus populaire fut surtout d'avoir pu changer la nuit en jour. Les êtres humains avaient toujours combattu l'obscurité quotidienne et paralysante de la nuit avec des feux de camp, des torches, des lampes à huile, puis des bougies ; mais durant un demi-million d'années, l'éclairage artificiel resta faible et vacillant.

Le XIXe siècle apporta quelques progrès dans ces méthodes d'éclairage vieillotes. On utilisa l'huile de baleine et le pétrole dans les lampes à huile, qui brillaient beaucoup plus et avaient un meilleur rendement. Le chimiste autrichien Karl Auer, baron von Welsbach, trouva que si on plaçait autour d'une flamme un cylindre de tissu imprégné de composés de thorium et de cérium, il émettait une lumière blanche brillante. Le *manchon Auer,* breveté en 1885, améliorait de façon appréciable la brillance de la lampe à pétrole.

Au début du XIXe siècle, l'inventeur écossais William Murdoch mit au point l'éclairage au gaz. Il envoya du gaz de charbon dans un tube et le fit sortir d'une buse où l'on pouvait l'enflammer. En 1802, il célébra la trêve conclue avec Napoléon en présentant un spectacle étonnant basé sur des éclairages au gaz ; vers 1803, son usine principale était entièrement éclairée au gaz. En 1807, certaines rues de Londres commencèrent à être éclairées au gaz, usage qui se répandit progressivement. Au fur et à mesure que s'écoulait le siècle, les grandes villes furent de mieux en mieux éclairées la nuit, ce qui diminua le taux de criminalité et augmenta la sécurité des citoyens.

Le chimiste américain Robert Hare découvrit qu'en dirigeant une flamme de gaz sur un bloc de chaux (*lime* en anglais), on obtenait une lumière blanche très brillante. On utilisa ce procédé (dit *limelight* en anglais) pour éclairer les scènes de théâtre avec plus de clarté qu'auparavant. Bien que ce procédé ne soit plus utilisé depuis longtemps, on parle toujours des « feux de la rampe ».

Toutes ces formes d'éclairage, du feu de camp au jet de gaz, utilisent des flammes. Il faut un dispositif pour allumer le combustible – que ce soit du bois, du charbon, du pétrole ou du gaz –, si une flamme n'existe

pas déjà à proximité. Avant le XIXe siècle, la méthode la moins laborieuse était d'utiliser un silex et de l'acier. En les frappant l'un contre l'autre, on obtenait des étincelles qui, avec un peu de chance, enflammaient un peu d'*étoupe* (matériau inflammable finement divisé), qui pouvait ensuite enflammer une bougie, et ainsi de suite.

Au début du XIXe siècle, les chimistes commencèrent à mettre au point des méthodes consistant à enduire l'extrémité d'un morceau de bois avec des produits chimiques susceptibles de s'enflammer quand on élevait la température. Le morceau de bois constituait une *allumette.* Le frottement élevait la température, et « frotter une allumette » sur une surface rugueuse produisait une flamme.

Les premières allumettes fumaient horriblement, dégageaient une mauvaise odeur et contenaient des produits chimiques toxiques. L'utilisation des allumettes devint réellement sûre en 1845, quand le chimiste autrichien Anton Ritter von Schrötter eut l'idée d'utiliser du phosphore rouge. Par la suite, on mit au point des *allumettes de sûreté* dans lesquelles le phosphore rouge était déposé sur une bande rugueuse sur la boîte d'allumettes, tandis que l'allumette elle-même contenait, dans sa tête, les autres produits chimiques nécessaires. Ni l'allumette, ni la bande rugueuse ne pouvaient s'enflammer seuls. Il fallait frotter l'allumette sur la bande pour qu'elle s'enflamme.

Il y a eu aussi un retour vers le silex et l'acier, avec des perfectionnements tout à fait remarquables. A la place du silex, on utilise le mischmetal, un alliage de métaux (au cérium principalement) qui émet, lorsqu'il est frotté contre une petite molette, des étincelles particulièrement chaudes. A la place de l'étoupe, on utilise un liquide aisément inflammable, l'*essence.* Le tout constitue le *briquet.*

L'ÉCLAIRAGE ÉLECTRIQUE

Les flammes, d'une espèce ou d'une autre, vacillent et créent un risque d'incendie. On éprouvait donc le besoin de quelque chose de totalement nouveau, et on avait remarqué depuis longtemps que l'électricité pouvait fournir de la lumière. Les bouteilles de Leyde fournissaient des étincelles quand on les déchargeait, les courants électriques portaient parfois à l'incandescence les fils qui les transportaient. Les deux phénomènes furent utilisés pour l'éclairage.

En 1805, Humphry Davy déclencha une décharge électrique dans l'espace d'air entre deux conducteurs. En maintenant le courant, ce qui rendait la décharge continue, il obtenait un *arc électrique.* Comme l'électricité devenait moins coûteuse, on se mit à utiliser des *lampes à arc* pour l'éclairage. Vers 1870, les rues de Paris et de quelques autres villes étaient équipées de ces lampes. Mais la lumière était violente, vacillante, à l'air libre ; elle constituait encore un risque d'incendie.

Il aurait été préférable de chauffer, à l'aide d'un courant électrique, un fil fin ou un filament jusqu'à ce qu'il devienne incandescent. Naturellement, il valait mieux que le filament soit porté à incandescence en l'absence d'oxygène qui, sinon, l'aurait fait brûler trop vite. Les premiers essais pour supprimer l'oxygène s'inspirèrent des méthodes déjà existantes pour évacuer l'air d'un récipient. Vers 1875, Crookes (au cours de son travail sur les rayons cathodiques ; voir chapitre 6) avait mis au point un

procédé pour produire un vide appréciable dans ce but, avec une vitesse suffisante et un faible coût. Néanmoins, les filaments utilisés restaient peu satisfaisants : ils cassaient trop aisément. En 1878, Thomas Edison, encore auréolé de son récent triomphe dû à l'invention du phonographe, annonça qu'il allait s'attaquer au problème. Il n'était âgé que de trente et un ans, mais sa réputation d'inventeur était telle que cette nouvelle entraîna une chute des actions des compagnies de gaz aux bourses de New York et de Londres.

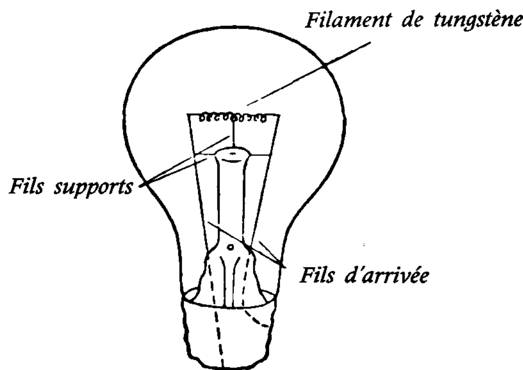
Après des centaines d'expériences et de déceptions, Edison trouva finalement le matériau qui servirait de filament – un fil de coton recouvert de noir de carbone. Le 21 octobre 1879, il alluma l'ampoule. Elle brûla quarante heures d'affilée. Pour la Saint-Sylvestre, Edison installa ses lampes, en guise de démonstration publique triomphante, dans la rue principale de Menlo Park, dans le New Jersey, où était situé son laboratoire. Il breveta très vite sa lampe et commença à la produire en masse.

Edison n'était pourtant pas le seul inventeur de la lampe à incandescence. Il y en eut au moins un autre qui pouvait à juste titre revendiquer l'invention. C'était l'Anglais Joseph Swan, qui présenta une lampe à filament de carbone à une rencontre de la Société chimique de Newcastle-on-Tyne le 18 décembre 1878. Toutefois il ne passa à la production de sa lampe qu'en 1881.

Edison s'attaqua alors au problème de l'approvisionnement des maisons en électricité, de manière à alimenter les lampes régulièrement et en quantité suffisante. Cette tâche lui demanda autant de créativité que l'invention de la lampe elle-même. On apporta plus tard deux importants perfectionnements à cette lampe. En 1910, William David Coolidge, de la General Electric Company, adopta le tungstène, métal réfractaire, comme constituant du filament (figure 9.7), et en 1913 Irving Langmuir mit dans l'ampoule un gaz peu réactif, l'azote, pour empêcher l'évaporation et la coupure du filament qui se produisaient dans le vide.

L'argon (dont l'usage date de 1920) remplit mieux encore cet objectif que l'azote, car il est complètement inerte. Le krypton, autre gaz inerte, est encore plus efficace ; il permet au filament de la lampe d'atteindre de plus hautes températures, donc de briller avec plus d'intensité sans diminution de sa durée de vie.

Figure 9.7. La lampe à incandescence.



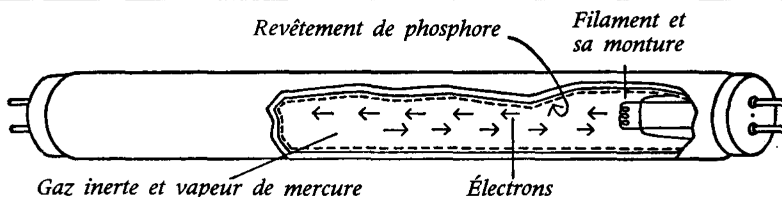
Pendant cinquante ans, le verre clair de l'ampoule transmit la lueur crue du filament incandescent, aussi difficile à regarder que le Soleil. Un ingénieur chimiste, Marvin Pipkin, mit au point une méthode pour dépolir l'intérieur du verre de l'ampoule (à l'extérieur, le dépoli aurait pu ramasser la poussière et diminuer la lumière émise). L'usage d'*ampoules dépolies* produisit une lumière douce et agréable.

L'avènement de l'éclairage électrique, en supprimant toutes les flammes non protégées servant à l'éclairage, laissait espérer qu'on pourrait éviter les incendies, si fréquents dans le passé. Hélas, il n'en a rien été car on utilise encore des flammes non protégées, ne serait-ce que dans les cheminées, les cuisinières à gaz, les radiateurs à gaz et à fuel. Il y a aussi les centaines de millions de fumeurs invétérés qui transportent avec eux des flammes non protégées sous la forme de cigarettes allumées et utilisent en outre fréquemment des briquets. Les dégâts et les pertes en vies humaines résultant des incendies dus aux cigarettes (feux de forêts et de broussailles aussi bien qu'incendies d'immeubles) sont difficiles à estimer.

Le filament chauffé à blanc de l'ampoule électrique (lumière *incandescente*, puisqu'elle est due à la chaleur dégagée à l'intérieur du filament par sa résistance à l'écoulement du courant électrique) n'est pas la seule façon de transformer l'électricité en lumière. Par exemple, les *lampes au néon* (introduites par le chimiste français Georges Claude en 1910) sont des tubes dans lesquels une décharge électrique excite les atomes du néon en leur faisant émettre une intense lueur rouge. La *lampe UV* contient de la vapeur de mercure qui émet, quand elle est excitée par une décharge, un rayonnement riche en ultraviolet ; on peut l'utiliser non seulement pour le bronzage mais aussi pour tuer les bactéries ou exciter de la fluorescence. Celle-ci, à son tour, conduit à l'*éclairage fluorescent,* introduit sous sa forme contemporaine en 1939 à la Foire mondiale de New York. C'est la lumière ultraviolette émise par la décharge dans la vapeur de mercure qui excite la fluorescence dans un *phosphore* revêtant l'intérieur du tube (figure 9.8). Comme cette lumière froide ne gaspille que peu d'énergie en chaleur, elle consomme moins d'énergie électrique.

Un tube fluorescent de 40 watts fournit autant de lumière et dégage bien moins de chaleur qu'une lampe incandescente de 150 watts. C'est pour cela que, depuis la Deuxième Guerre mondiale, on a assisté en matière d'éclairage à un renversement massif vers le fluorescent. Les premiers tubes utilisaient des sels de béryllium comme phosphores, ce qui a entraîné des cas d'empoisonnement sérieux *(bérylliose)* dus à l'inhalation de poussières contenant ces sels ou à des coupures par des débris de tubes

Figure 9.8. La lampe fluorescente. Une décharge dans les électrons émis par le filament excite la vapeur de mercure dans le tube, produisant un rayonnement ultraviolet. L'ultraviolet fait s'illuminer le phosphore.



cassés. A partir de 1949, on a utilisé des phosphores beaucoup moins dangereux.

Le perfectionnement le plus prometteur semble bien être une méthode de conversion directe de l'électricité en lumière sans création préalable de lumière ultraviolette. En 1936, le physicien français Georges Destriau a découvert qu'un courant alternatif intense pouvait faire émettre une lueur à du phosphore. Aujourd'hui, les ingénieurs répartissent le phosphore dans du plastique ou du verre et utilisent ce phénomène, appelé *électroluminescence*, pour mettre au point des panneaux lumineux. Ainsi, un mur ou un plafond luminescent pourrait éclairer une pièce, la faisant baigner dans une douce lueur colorée. Le rendement de l'électroluminescence est cependant trop faible pour lui permettre de concurrencer les autres procédés d'éclairage électrique.

LA PHOTOGRAPHIE

Aucune invention mettant en jeu la lumière n'a probablement apporté à l'humanité autant de plaisir que la photographie. On avait observé que la lumière passant à travers un trou d'épingle dans une petite chambre noire (*camera obscura* en latin) donnait une faible image inversée de la scène extérieure à la chambre noire. C'est un alchimiste italien, Giambattista Della Porta, qui construisit vers 1550 ce dispositif, appelé *chambre noire* ou *sténopé.*

La quantité de lumière qui pénètre dans une chambre noire est très faible. Si en revanche on substitue une lentille au trou d'épingle, on peut rassembler une quantité considérable de lumière dans un plan dit plan focal, et l'image est beaucoup plus brillante. Lorsque l'on eut obtenu ce résultat, il fallait trouver une réaction chimique sensible à la lumière. De nombreux expérimentateurs travaillèrent dans ce but, en particulier les Français Joseph Nicéphore Niepce et Louis Jacques Mande Daguerre, ainsi que l'Anglais William Henry Fox Talbot. Niepce essaya de faire noircir du chlorure d'argent à la lumière du jour selon une image donnée. Il obtint la première photographie en 1822, mais un temps de pose de huit heures lui était nécessaire.

Daguerre s'associa à Niepce peu avant la mort de ce dernier pour améliorer le procédé. Après le noircissement au soleil de sels d'argent, il dissolvait les sels non modifiés dans du thiosulfate de sodium, procédé suggéré par l'homme de science John Herschel (fils de William Herschel). Vers 1839, Daguerre arriva à produire les *daguerréotypes*, les premières photographies réalisables, avec des temps de pose n'excédant pas vingt minutes.

Talbot améliora encore le procédé, en obtenant des *négatifs* dans lesquels les régions éclairées de l'image sont assombries, de sorte que le noir devient blanc et vice versa. En 1844, Talbot publia le premier livre illustré de photographies.

La photographie vit son utilité démontrée pour la documentation quand, vers 1850, Talbot photographia des scènes de la guerre de Crimée et quand, dix ans plus tard, le photographe américain Matthew Brady, avec un équipement que nous considérerions aujourd'hui comme très primitif, prit des photographies de la guerre de Sécession en Amérique du Nord.

Pendant près d'un demi-siècle, la *plaque humide* fut utilisée en photographie. Elle consistait en une plaque de verre, sur laquelle

on répandait une émulsion de produits chimiques qu'on devait préparer sur-le-champ. Tant qu'il ne fut pas possible d'échapper à cette contrainte, seuls des professionnels bien entraînés purent prendre des photographies.

En 1878, un inventeur américain, George Eastman, découvrit comment mélanger l'émulsion à de la gélatine, comment l'étaler sur la plaque et la laisser sécher de façon à obtenir un gel solide qui pouvait rester actif longtemps. En 1884, il fit breveter la *pellicule* photographique (*film* en anglais), dans laquelle le gel était étalé sur du papier. En 1889, il remplaça le papier par du celluloïd. En 1888, il inventa le Kodak, un appareil qui prenait des photos sur la seule pression d'un bouton. Il fallait, bien sûr, donner ensuite le film exposé à développer. La photographie devint alors un passe-temps populaire. Comme on introduisait sur le marché des émulsions toujours plus sensibles, les vues pouvaient être prises en un clin d'œil et il n'y avait plus besoin de garder la pose, qui donnait des expressions figées et artificielles.

On pensait que l'on ne pouvait pas simplifier davantage le procédé ; toutefois, en 1947, l'inventeur américain Edwin Herbert Land a mis au point un appareil photo muni d'un double rouleau de pellicule, l'un de film négatif ordinaire, l'autre de papier positif, avec des capsules scellées de produits chimiques entre les deux. Les produits chimiques sont libérés au moment voulu et développent automatiquement l'épreuve positive. On a la photo en main quelques minutes seulement après l'avoir prise.

Durant tout le XIXᵉ siècle, les photographies étaient en noir et blanc, sans couleurs. Au début du XXᵉ siècle, le physicien français d'origine luxembourgeoise Gabriel Lippmann mit au point un procédé de photographie en couleurs qui lui valut le prix Nobel de physique en 1908. Cela s'avéra être un faux départ, et il fallut attendre 1936 pour que soit mis au point un procédé exploitable de photographie en couleurs. Ce deuxième essai, couronné de succès, était basé sur l'observation faite en 1855 par Maxwell et von Helmholtz que n'importe quelle couleur du spectre pouvait s'obtenir par combinaison de rouge, vert et bleu. La pellicule couleurs basée sur ce principe est composée de trois couches d'émulsion – l'une sensible à la composante rouge, la deuxième à la composante verte et la troisième à la composante bleue de l'image. Trois photos séparées mais superposées se forment, chacune reproduisant l'intensité de la lumière dans son domaine de spectre sous la forme d'une image noir et blanc. La pellicule est alors développée en trois étapes successives, en utilisant des colorants rouge, bleu et vert pour déposer les couleurs appropriées sur le négatif. Chaque point de l'image finale est une combinaison spécifique de rouge, de vert et de bleu, et le cerveau interprète ces combinaisons pour reconstituer toutes les nuances de couleurs.

En 1959, Land a présenté une nouvelle théorie de la vision des couleurs. Le cerveau, selon lui, n'a pas besoin d'une combinaison de trois couleurs pour créer l'impression de la couleur. Tout ce qu'il lui faut, c'est deux longueurs d'onde différentes, ou deux groupes de longueurs d'onde, séparés l'un de l'autre par un intervalle minimal de longueur d'onde. Par exemple, l'un des groupes de longueurs d'onde peut être constitué du spectre tout entier, c'est-à-dire de la lumière blanche. Comme la longueur d'onde moyenne de la lumière blanche est dans la région jaune-vert, celle-ci peut

servir de longueur d'onde « courte ». Dans ces conditions, une image reproduite par la combinaison d'une image en lumière blanche et d'une image en lumière rouge (servant de longueur d'onde longue) reproduit toutes les couleurs. Land a aussi obtenu des images en couleurs avec des lumières filtrées verte et rouge et d'autres combinaisons de deux longueurs d'onde.

L'invention du cinéma trouve ses sources dans une observation faite pour la première fois par le médecin anglais Peter Mark Roget en 1824. Il remarqua que l'œil forme une image persistante, qui dure une fraction appréciable de seconde. Après la découverte de la photographie, de nombreux expérimentateurs, particulièrement en France, utilisèrent cette propriété pour créer l'illusion du mouvement en montrant une série d'images en succession rapide. Chacun connaît le jeu constitué d'une suite de cartes portant chacune une image ; quand on découvre tour à tour chaque carte rapidement à l'aide du pouce, le spectateur observe une image animée. De même, si on projette rapidement sur un écran une série d'images, chacune légèrement différente de celle qui la précède, à des intervalles d'environ 1/16 de seconde, la persistance des images successives sur la rétine de l'œil les fait s'enchaîner les unes aux autres, donnant ainsi l'impression d'un mouvement continu.

Ce fut Édison qui produisit le premier *film cinématographique*. Il photographia une série d'images sur une bande de film, puis fit passer celui-ci dans un projecteur qui les rendit chacune visible successivement au moyen d'un éclair de lumière. Le premier film fut présenté au public en 1894* ; en 1914, les cinémas américains projetèrent le premier film cinématographique de long métrage, *La naissance d'une nation.*

Au film muet, on ajouta en 1927 une *piste sonore*. La piste sonore utilise aussi la lumière : la configuration des ondes sonores de la musique et des voix des acteurs est convertie, au moyen d'un microphone, en un courant électrique variable ; ce courant excite une lampe qui est photographiée en synchronisme avec l'action du film. Quand on projette le film muni d'une telle piste sur l'un de ses côtés, la variation d'intensité lumineuse à travers la piste est convertie en courant électrique par une *cellule photo-électrique* utilisant l'effet du même nom, et ce courant est enfin reconverti en son.

Dans les deux ans suivant le premier *film parlant, Le chanteur de jazz,* les films muets appartenaient au passé, comme le vaudeville. A la fin des années trente apparut la couleur. Les années cinquante virent, en outre, se développer les techniques des écrans élargis, et connurent même un engouement de faible durée pour les effets de relief (3D), nécessitant la projection de deux images sur un même écran. En portant des lunettes à verres polarisés, le spectateur percevait une image distincte pour chaque œil, ce qui produisait l'effet stéréoscopique recherché.

---

* Sans vouloir entrer dans des querelles stériles de paternité d'invention, rappelons néanmoins qu'en 1894 il s'agissait du *kinétoscope* d'Édison, appareil où l'on observe des vues animées avec une paire d'oculaires. Au cours de l'année qui suivit, de nombreux pionniers en Europe et aux États-Unis mirent au point diverses méthodes pour la projection des images animées sur écran. De tous ces systèmes, celui sans conteste le plus achevé fut celui des frères Auguste et Louis Lumière. Breveté en février 1895 sous le nom de *cinématographe*, il fut présenté triomphalement au public parisien en décembre de la même année. *(N.D.T.)*

Les moteurs à combustion interne

Le *pétrole lampant*, ou kérosène, avait précédé l'électricité dans le domaine de l'éclairage. Un autre produit de la distillation du pétrole brut, *l'essence*, allait bientôt induire un autre développement technique, développement qui révolutionna la vie moderne aussi profondément que l'introduction de l'électricité. Il s'agit du *moteur à combustion interne*, ainsi nommé parce que, dans un tel moteur, le carburant est brûlé dans le cylindre de façon que les gaz formés poussent directement le piston. Les machines à vapeur ordinaires sont des *moteurs à combustion externe*, car le carburant est brûlé à l'extérieur du cylindre, où l'on introduit la vapeur déjà formée.

L'AUTOMOBILE

Ce dispositif compact, avec de petites explosions circonscrites au cylindre, permit de fournir de l'énergie motrice à de petits véhicules dans des conditions auxquelles l'encombrante machine à vapeur n'était pas adaptée. Des « voitures sans chevaux », actionnées par la vapeur, furent réalisées dès 1786. L'une d'elles fut construite par William Murdoch, qui devait plus tard inventer l'éclairage au gaz. Un siècle plus tard, l'inventeur américain Francis Edgar Stanley inventa la célèbre *Stanley Steamer*, qui fut un temps en compétition avec les premières voitures équipées de moteurs à combustion interne. L'avenir devait appartenir à ces derniers.

En fait, on construisit quelques moteurs à combustion interne dès le début du XIXe siècle, avant que le pétrole ne devienne un produit d'usage commun. On utilisait alors pour ces moteurs des vapeurs d'essence de térébenthine ou de l'hydrogène. Mais ce ne fut qu'avec l'essence, le seul liquide se vaporisant facilement, à la fois combustible et disponible en grande quantité, qu'un tel moteur put devenir autre chose qu'une curiosité.

Le premier moteur à combustion interne utilisable fut construit en 1860 par l'inventeur français Étienne Lenoir, qui le monta dans une petite carriole, la première « voiture sans chevaux » munie d'un tel moteur. En 1876, le technicien allemand Nikolaus August Otto, qui avait entendu parler du moteur Lenoir, construisit un moteur *à quatre temps* (figure 9.9). Dans le premier temps, un piston étroitement ajusté dans un cylindre est tiré vers l'extérieur, de façon à aspirer un mélange d'essence et d'air dans le cylindre vide. Puis le piston est repoussé dans le cylindre pour comprimer les gaz. Au point de compression maximale, les gaz sont enflammés et explosent. L'explosion repousse le piston, et c'est ce mouvement qui fait tourner le moteur. Il entraîne une roue qui fait rentrer de nouveau le piston pour expulser du cylindre les résidus de combustion ou *gaz d'échappement* – quatrième et dernier temps du cycle. La roue fait de nouveau sortir le piston ; cela constitue le début du cycle suivant.

Un ingénieur écossais, Dugald Clerk, trouva presque immédiatement un perfectionnement au moteur à quatre temps. Il ajouta un deuxième cylindre, disposé de façon que son piston fût entraîné par l'explosion au moment où l'autre était dans le temps d'admission : cette disposition rendait le débit d'énergie beaucoup plus régulier. Plus tard, l'addition d'autres cylindres (huit est un nombre courant aujourd'hui) augmenta la régularité et la puissance de ce *moteur à explosion*.

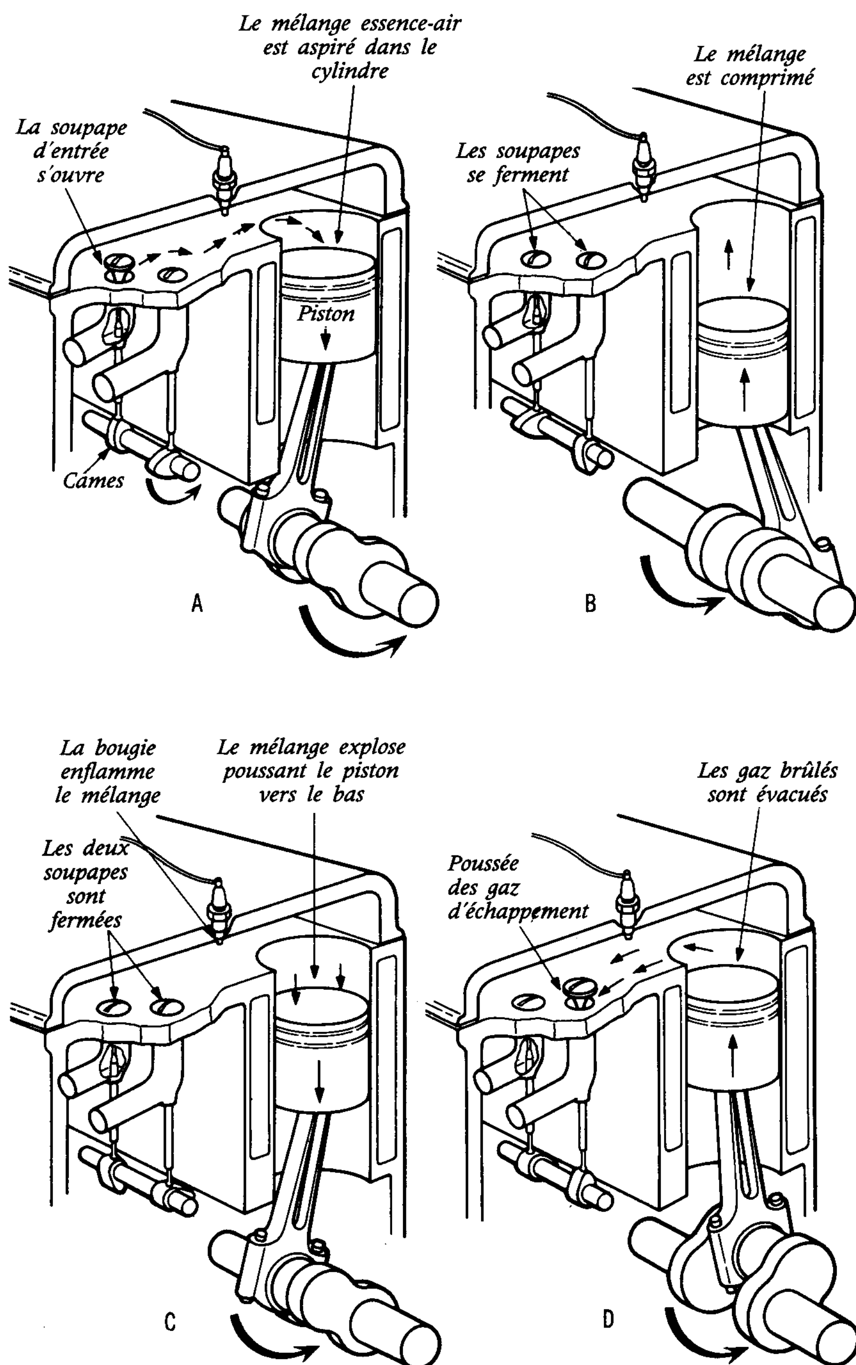


Figure 9.9. Le moteur à quatre temps de Nikolaus Otto, construit en 1876.

Un tel moteur s'avérait essentiel à la réalisation pratique des automobiles, mais des inventions auxiliaires s'imposaient aussi. L'allumage du mélange d'essence et d'air au bon moment posait un problème difficile. On utilisa pour cela toutes sortes de dispositifs ingénieux ; vers 1923 il devint courant de s'en remettre à un système électrique. L'alimentation provient d'une *batterie* d'accus, qui, comme une batterie de piles, fournit l'électricité résultant d'une réaction chimique. Mais ici on peut recharger les accus en les faisant traverser par un courant électrique de sens opposé à celui de leur décharge ; ce courant renverse le sens de la réaction chimique et les produits chimiques peuvent alors refournir de l'électricité. Le courant inverse est apporté par un petit générateur qu'entraîne le moteur.

Le type le plus commun de batterie d'accumulateurs a une alternance de plaques de plomb et d'oxyde de plomb, trempant dans une solution relativement concentrée d'acide sulfurique. Il a été inventé par le physicien français Gaston Planté en 1859 et mis sous sa forme moderne par l'ingénieur électricien américain Charles Francis Brush en 1881. On a inventé depuis des accumulateurs plus robustes et plus compacts – par exemple un accumulateur nickel-fer réalisé par Edison vers 1905 – mais d'un point de vue économique, aucun ne peut rivaliser avec la batterie au plomb.

La tension électrique fournie par la batterie est emmagasinée dans le champ magnétique d'un transformateur appelé *bobine d'induction*, et la chute brutale de ce champ fournit la pointe de tension produisant l'étincelle entre les électrodes des bougies d'allumage qui nous sont familières.

Une fois qu'un moteur à combustion interne a démarré, l'inertie le maintient en mouvement entre les temps moteurs. Mais il faut lui fournir de l'énergie pour le faire démarrer. Au début, cela se faisait à la main (avec la manivelle de l'automobile), et on fait encore démarrer à la main les moteurs hors-bord et les tondeuses à gazon en tirant sur une ficelle. La manivelle de l'automobile, elle, exigeait un bras bien musclé. Quand le moteur commençait à tourner, il était assez fréquent de voir la manivelle échapper à la main qui la tenait, tourner et casser le bras de l'opérateur. En 1912, l'Américain Charles Franklin Kettering inventa un *démarreur* qui permettait d'en finir avec la manivelle. Le démarreur est alimenté par la batterie qui fournit l'énergie nécessaire aux premiers tours du moteur.

Ce furent deux ingénieurs allemands, Gottlieb Daimler et Carl Benz, qui construisirent indépendamment en 1885 les premières automobiles d'utilisation pratique. Mais ce qui fit de l'automobile un moyen de transport usuel, ce fut l'invention de la *production en série*.

Le premier à mettre en œuvre cette technique fut Eli Whitney, qui mérite plus de renommée pour cela que pour sa bien plus célèbre invention de la machine à égrener le coton. En 1789, Whitney signa un contrat avec le gouvernement fédéral américain pour la fourniture de fusils à l'armée. Jusqu'à cette époque, les fusils avaient été fabriqués un par un, chacun avec ses propres pièces.

Whitney conçut des pièces détachées standard, de façon qu'une pièce donnée puisse s'adapter à n'importe quel fusil. Cette innovation simple et unique – un standard d'usinage, des pièces détachées interchangeables pour un type donné de produit – fut peut-être aussi importante que n'importe quel autre facteur dans la création de l'industrie moderne de production

de masse. Quand de puissantes machines furent disponibles, elles permirent de fabriquer des pièces standardisées en nombre pratiquement illimité.

L'ingénieur américain Henri Ford, le premier, alla jusqu'au bout de ce concept. Il avait construit sa première automobile (une deux-cylindres) en 1892, puis était parti travailler pour la Detroit Automobile Company en 1899 en qualité d'ingénieur en chef. La compagnie voulait produire des voitures sur mesure, mais Ford avait d'autres desseins. Il démissionna en 1902 pour produire des automobiles à son compte, en quantité. En 1909, il commença à produire le modèle T, et vers 1913, à le fabriquer selon la méthode de Whitney : une voiture après l'autre, chacune identique à la précédente, et toutes avec les mêmes pièces détachées.

Ford comprit qu'il pouvait accélérer la production en utilisant des travailleurs humains comme on utilisait des machines, à faire le même petit travail de façon répétitive sans interruption. L'Américain Samuel Colt (qui avait inventé le revolver ou « six-coups ») avait posé les premiers jalons de la méthode en 1847, et le fabricant d'automobiles Ransom E. Olds avait appliqué le système au moteur d'automobile en 1900. Cependant, Olds perdit son support financier, et il échut donc à Ford de mener à bien le projet. Ford conçut la *chaîne,* avec des travailleurs ajoutant des pièces à l'assemblage alors que celui-ci passait devant eux sur des bandes transporteuses, jusqu'à ce que la voiture terminée quitte la chaîne en roulant. Deux progrès économiques résultèrent de ce système : de hauts salaires pour les ouvriers, et des voitures qu'on pouvait vendre à des prix étonnamment bas.

Vers 1913, Ford construisait mille modèles T par jour. Lorsque la chaîne fut arrêtée en 1927, on avait produit quinze millions de voitures et le prix était tombé à 290 dollars. L'engouement pour le changement de véhicule chaque année l'emporta alors, et Ford fut forcé de rejoindre les tenants de la variété et de la nouveauté superficielle, qui fit augmenter énormément le prix des automobiles et perdre aux automobilistes une grande part des avantages de la production en série.

En 1892, l'ingénieur mécanicien allemand Rudolf Diesel introduisit une modification au moteur à combustion interne, qui était simple et économique en carburant. Il portait le mélange carburant-air à haute pression, de sorte que la chaleur de la compression seule suffisait à l'enflammer. Le *moteur diesel* rend possible l'utilisation de fractions de distillation du pétrole à plus haut point d'ébullition qui ne cliquètent pas. A cause de la plus haute compression utilisée, le moteur doit être plus solide. Il est donc bien plus lourd que le moteur à essence. Une fois qu'un système adéquat d'injection du combustible fut mis au point dans les années vingt, il commença à s'imposer pour les camions, les tracteurs, les autocars, les bateaux ; il est maintenant le roi incontesté du transport lourd.

Des améliorations de l'essence augmentèrent encore le rendement du moteur à combustion interne. L'essence est un mélange complexe de molécules faites d'atomes de carbone et d'hydrogène (*hydrocarbures*), dont certaines brûlent plus vite que d'autres. Une vitesse de combustion trop rapide est à éviter, car alors le mélange essence-air explose en trop d'endroits à la fois, produisant le *cliquetis.* Une vitesse plus faible de combustion produit une dilatation égale de la vapeur, qui pousse le piston régulièrement et efficacement.

Le taux de cliquetis d'une essence donnée est mesuré par son indice d'octane, par comparaison avec le cliquetis d'un hydrocarbure appelé *iso-octane,* qui produit très peu de cliquetis, mélangé à de l'*heptane normal,*

qui en produit beaucoup. L'un des buts principaux du raffinage de l'essence est de fabriquer un mélange d'hydrocarbures avec un indice d'octane élevé.

On a conçu des moteurs d'automobiles avec des *taux de compression* de plus en plus élevés ; cela signifie que le mélange essence-air est comprimé à de plus grandes densités avant l'allumage. Cette compression permet d'extraire plus d'énergie de l'essence mais encourage aussi le cliquetis, de sorte qu'il faut fabriquer de l'essence à indice d'octane de plus en plus élevé.

Pour simplifier la tâche, on utilise des produits chimiques qui, quand on les ajoute en petite quantité à l'essence, réduisent le cliquetis. Le plus efficace de ces *composés anti-cliquetis* est le *plomb tétraéthyle,* un composé du plomb dont les propriétés avaient été remarquées par le chimiste américain Thomas Midgley et qu'on utilisa pour la première fois en 1925. L'essence qui en contient est appelée *supercarburant.* S'il n'y avait que du plomb tétraéthyle, les oxydes de plomb formés durant la combustion de l'essence encrasseraient et détruiraient le moteur. C'est pour cette raison qu'on ajoute du bromure d'éthylène. L'atome de plomb du plomb tétraéthyle se combine avec l'atome de brome du bromure d'éthylène pour former du bromure de plomb qui, à la température de combustion de l'essence, se vaporise et est expulsé avec les gaz d'échappement.

Les carburants diesel sont testés, en ce qui concerne le temps s'écoulant entre compression et allumage (pour éviter un temps trop long), par comparaison avec un hydrocarbure appelé *cétane,* dont la molécule contient seize atomes de carbone, alors que celle d'iso-octane n'en contient que huit. Pour les carburants diesel, on parle donc d'*indice de cétane.*

Les perfectionnements continuèrent. En 1923, ce furent les pneus basse pression, au début des années cinquante les pneus sans chambre, qui rendirent les éclatements moins fréquents. Dans les années quarante, les voitures eurent l'air conditionné aux États-Unis, puis les transmissions automatiques furent introduites et les changements de vitesse commencèrent à être moins utilisés. Vers les années cinquante apparurent la direction et les freins assistés. L'automobile est devenue une part si intrinsèque de la vie moderne qu'en dépit de l'augmentation du coût de l'essence et de la pollution atmosphérique, il semble impossible, à l'exception d'une catastrophe détruisant l'humanité, qu'elle disparaisse.

L'AVION

L'autobus et le poids lourd constituent des versions géantes de l'automobile ; le pétrole a remplacé le charbon sur les gros bateaux, mais le plus grand succès remporté par le moteur à combustion interne a eu lieu dans les airs. Vers 1890, les hommes avaient réalisé le vieux rêve – plus vieux que Dédale et Icare – de voler avec des ailes. Le planeur était devenu un sport d'amateurs enthousiastes. En 1853, l'inventeur anglais George Cayley avait construit le premier *planeur* pouvant porter un homme. L'« homme » qu'il transportait n'était cependant qu'un jeune garçon. Le premier à avoir véritablement pratiqué ce sport, l'ingénieur allemand Otto Lilienthal, se tua en 1896. Mais on eut très vite envie d'entreprendre des vols à moteur ; le planeur, lui, est resté populaire en tant que sport.

Le physicien et astronome américain Samuel Pierpont Langley essaya, en 1902 et 1903, de faire voler un planeur propulsé par un moteur à combustion interne et fut à deux doigts de réussir. S'il avait encore eu

de l'argent, il aurait volé lors de l'essai suivant. Cet honneur fut réservé aux frères Orville et Wilbur Wright, fabricants de bicyclettes qui pratiquaient le planeur comme passe-temps.

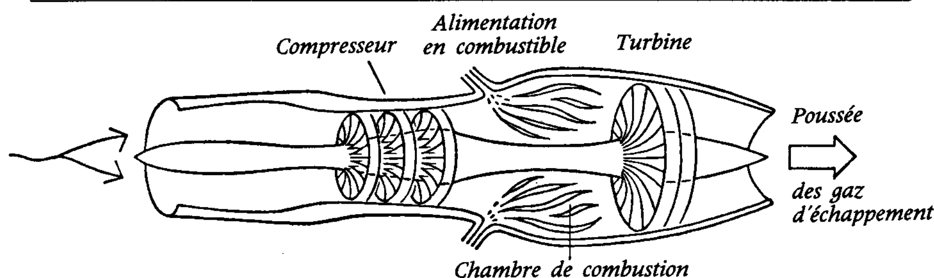
Le 17 décembre 1903, à Kitty Hawk, en Caroline du Nord, les frères Wright quittèrent le sol dans un planeur propulsé par un moteur et restèrent en l'air 59 secondes, parcourant 16 mètres. C'était le premier vol en avion de l'histoire, et il passa quasi inaperçu aux yeux du monde entier.

Le public manifesta davantage d'intérêt après que les Wright eurent effectué des vols de plus de 30 kilomètres et quand, en 1909, l'ingénieur français Louis Blériot traversa la Manche en aéroplane. Les batailles aériennes et les exploits de la Première Guerre mondiale stimulèrent les imaginations ; les *biplans* de cette époque, aux deux ailes précairement reliées par des haubans et des montants, devinrent familiers aux spectateurs de cinéma de la génération d'après-guerre. L'ingénieur allemand Hugo Junkers réalisa un *monoplan* juste après la guerre, et l'aile simple épaisse, sans montants, l'emporta définitivement. (En 1939, l'ingénieur américain d'origine russe Igor Ivan Sikorsky construisit un avion multimoteur et conçut le premier *hélicoptère*, avion aux pales rotatives sustentatrices qui peut décoller et atterrir à la verticale, et même faire du surplace.)

Mais, au début des années vingt, l'avion restait plus ou moins une curiosité – simple instrument de guerre nouveau et plus horrible encore, jouet pour aviateurs casse-cou et amateurs de suspense. L'aviation ne fut reconnue que lorsque, en 1927, Charles Auguste Lindbergh vola sans escale de New York à Paris. Le monde s'enthousiasma pour cet exploit et l'on commença à concevoir des avions plus gros et plus sûrs.

Deux grandes innovations ont été apportées au moteur d'avion depuis qu'il sert de moyen de transport. La première est l'adoption de la turbine à gaz (figure 9.10). Dans ce moteur, les gaz chauds en train de se dilater entraînent une turbine par pression contre ses ailettes, au lieu de déplacer des pistons dans des cylindres. Le moteur est simple, moins coûteux à l'entretien, et moins sujet aux pannes ; il a seulement fallu attendre, pour qu'il devienne d'usage commun, la mise au point d'alliages qui puissent supporter les hautes températures des gaz chauds. On conçut de tels alliages dès 1939. Peu après, les *avions à turbo-propulseurs,* utilisant une turbine à gaz pour entraîner des hélices, se répandirent de plus en plus.

Figure 9.10. Un turboréacteur. De l'air est aspiré, comprimé et mélangé avec du combustible ; le mélange est enflammé dans la chambre de combustion. La dilatation des gaz entraîne une turbine et produit une poussée.



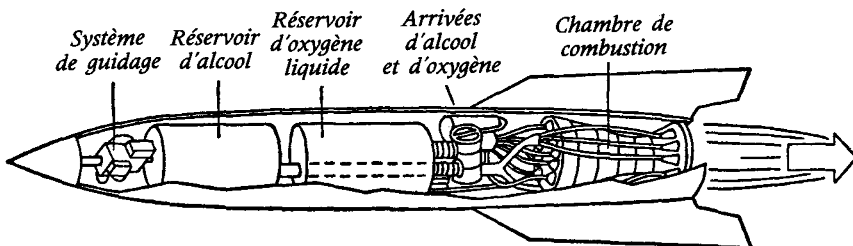
Mais ils furent vite surpassés, au moins pour les longs courriers, par le deuxième perfectionnement, l'avion à *turboréacteur.* En principe, la force d'entraînement est dans ce cas la même que celle qui fait se déplacer un ballon de baudruche gonflé dont on relâche l'orifice pour permettre à l'air de s'échapper. C'est l'action et la réaction : le mouvement de l'air en expansion, qui s'échappe dans une direction, entraîne un mouvement équivalent ou poussée dans la direction opposée – exactement comme le mouvement vers l'avant de la balle dans un canon de fusil fait reculer ce dernier vers l'arrière. Dans le turboréacteur, la combustion du carburant produit des gaz chauds à haute pression qui propulsent l'avion vers l'avant avec une grande force au fur et à mesure de leur violente éjection vers l'arrière par la tuyère de sortie. Une fusée est propulsée de la même façon, à ceci près qu'elle emporte son propre oxygène pour brûler le combustible (figure 9.11).

Les brevets pour la *propulsion à réaction* furent déposés par un ingénieur français, René Lorin, dès 1913 ; mais, à cette époque, c'était une idée complètement irréalisable. La propulsion par réaction n'est économique qu'à des vitesses de plus de 800 km/h. En 1939, un Anglais, Franck Whittle, fit voler un avion à réaction raisonnablement pratique d'utilisation ; en janvier 1944, des avions à réaction furent utilisés par l'Angleterre et les États-Unis contre les *bombes volantes* allemandes, les V1, un avion sans pilote transportant des explosifs dans son nez.

Après la Deuxième Guerre mondiale, on mit au point des avions à réaction militaires qui approchaient de la vitesse du son. La vitesse du son dépend de l'élasticité naturelle des molécules d'air, de leur aptitude à faire des va-et-vient. Quand l'avion approche cette vitesse, les molécules d'air ne peuvent plus s'écarter sur son passage et sont comprimées en avant de l'avion, qui subit donc toute une série de contraintes et d'efforts. On parla du *mur du son* comme si c'était là quelque chose de physique qui ne pouvait être approché sans risques énormes. Cependant, des essais en soufflerie conduisirent à des profils plus aérodynamiques, et le 14 octobre 1947, un avion-fusée américain, le X1, piloté par Charles Elwood Yeager, « franchit le mur du son ». Pour la première fois dans l'histoire, un être humain dépassait la vitesse du son. Les victoires aériennes de la guerre de Corée, au début des années cinquante, furent le fait d'avions à réaction qui volaient à des vitesses telles qu'ils ne pouvaient pratiquement pas être atteints et descendus.

Le rapport de la vitesse d'un objet à celle du son (qui vaut 1 194,1 km/h à 0° C) dans le milieu à travers lequel l'objet se déplace est le *nombre de Mach,* d'après le physicien autrichien Ernst Mach, qui étudia le premier, au milieu du XIXe siècle, les conséquences du mouvement à de telles

Figure 9.11. Une fusée à combustible liquide.



vitesses. Vers 1960, la vitesse des avions dépassait Mach 5. Sur l'avion-fusée expérimental X15, les fusées le propulsaient si haut durant de courtes périodes de temps que ses pilotes se qualifiaient d'astronautes. Les avions militaires volent à des vitesses plus basses et les avions commerciaux à des vitesses encore inférieures.

Un avion volant à *vitesse supersonique* (au-dessus de Mach 1) transporte ses ondes sonores en avant de lui puisqu'il se déplace plus vite que les ondes sonores seules ne le pourraient. S'il est assez proche du sol, le cône des ondes sonores compressées peut intercepter le sol avec un *bang sonore.* (Le claquement d'un fouet est un bang miniature, puisque convenablement manipulé, le bout du fouet peut atteindre une vitesse supersonique.)

Le vol supersonique commercial a été inauguré par l'avion franco-britannique *Concorde,* qui traverse l'Atlantique en trois heures, volant à deux fois la vitesse du son. Une version américaine d'un tel TSS *(transport supersonique),* a été abandonné à l'état de projet en 1971, en raison de craintes touchant à l'augmentation excessive du bruit au-dessus des aéroports et des risques d'atteinte à l'environnement. Quelques personnes ont alors remarqué que c'était la première fois qu'un progrès technologique réalisable avait été arrêté en raison de son inopportunité. C'était la première fois que des êtres humains avaient dit : « Nous pouvons le faire, mais il vaut mieux ne pas le faire. »

Pour conclure, c'est peut-être tout aussi bien, car les gains ne semblent pas justifier la dépense. Le *Concorde* a été un échec économique, et le programme de TSS soviétique n'a pas survécu à l'écrasement au sol de l'un de ses avions lors d'une démonstration à Paris en 1973.

L'électronique

LA RADIO

En 1888, Heinrich Hertz mena à bien les célèbres expériences qui mirent en évidence les ondes radio, prédites vingt ans plus tôt par James Clerk Maxwell (voir le chapitre 8). Hertz établit un courant alternatif de haute tension qui passait dans deux boules de métal séparées par un petit intervalle d'air. Chaque fois que le potentiel atteignait son maximum dans une direction ou dans l'autre, une étincelle jaillissait entre les deux boules. Dans ces conditions, les équations de Maxwell prédisaient l'émission d'un rayonnement électromagnétique. Hertz utilisa, pour détecter cette énergie, un récepteur constitué d'une simple boucle de fil munie d'une coupure en un point. De même que le courant donnait naissance à un rayonnement dans le premier appareil, le rayonnement devait donner naissance à un courant dans la boucle. En effet, Hertz put détecter de petites étincelles traversant sa boucle réceptrice, même lorsque celle-ci était placée à l'autre bout de la pièce, loin de l'appareil émetteur. L'énergie était donc transmise à travers l'espace.

En déplaçant sa boucle détectrice en divers endroits de la pièce, Hertz put déterminer la forme des ondes. Là où les étincelles étaient brillantes, on avait un maximum ou un minimum. Là où elles n'apparaissaient pas du tout, on se trouvait entre maximum et minimum. Il pouvait donc calculer la longueur d'onde du rayonnement. Il trouva que ces ondes étaient incomparablement plus longues que celles de la lumière.

Au cours de la décade suivante, il apparut à nombre de gens que les *ondes hertziennes* pouvaient servir à transmettre des messages d'un endroit à un autre, car les ondes étaient assez longues pour contourner les obstacles. En 1890, le physicien français Édouard Branly réalisa un récepteur perfectionné en remplaçant la boucle de fil par un tube en verre rempli de limaille métallique, auquel étaient reliés des conducteurs et une pile. La limaille ne conduisait pas le courant de la pile, à moins qu'une haute tension alternative n'y fût induite, ce que faisaient justement les ondes hertziennes. Avec ce récepteur, il fut capable de détecter des ondes hertziennes à une distance de 140 mètres. Puis le physicien anglais Oliver Joseph Lodge (qui devait s'illustrer plus tard comme champion du spiritisme) modifia ce dispositif ; il réussit à détecter des signaux à une distance de 800 mètres et à envoyer des messages en code morse.

L'inventeur italien Guglielmo Marconi découvrit qu'il pouvait améliorer les choses en connectant un pôle du générateur et du récepteur à la terre, et en reliant l'autre à un fil appelé plus tard *antenne* (parce qu'il ressemblait, je suppose, à une antenne d'insecte). En utilisant de puissants générateurs, Marconi parvint à envoyer des signaux à une distance de 145 kilomètres en 1896, à leur faire traverser la Manche en 1898 et l'Atlantique en 1901. Ainsi était née la *télégraphie sans fil, T.S.F.* ou *radio.*

Marconi travailla sur un système destiné à supprimer les *parasites* dus à d'autres sources et à accorder le récepteur à la longueur d'onde générée par l'émetteur. Pour ses inventions, Marconi partagea le prix Nobel de physique en 1909 avec le physicien allemand Karl Ferdinand Braun, qui avait contribué aussi au développement de la radio en montrant que certains cristaux ont la propriété de ne laisser passer le courant que dans une seule direction. C'est ainsi que du courant alternatif pouvait être converti en courant continu, ce dont les radios avaient justement besoin. Les cristaux tendaient à être instables ; dans les années dix, les gens se penchaient sur leur *appareil à cristal de galène* pour recevoir des signaux.

Le physicien américain Reginald Aubrey Fessenden mit au point un générateur inédit de courants alternatifs à haute fréquence (rompant avec le système à éclateur à étincelles) et développa un système de *modulation* de l'onde radio qui permet à cette dernière de transporter une structure imitant les ondes sonores. Ce qui était modulé, c'était l'amplitude (ou hauteur) des ondes ; on appela donc cela *modulation d'amplitude,* d'où le nom de *radio MA.* Au soir de Noël 1906, la musique et la parole sortirent pour la première fois d'un récepteur de radio.

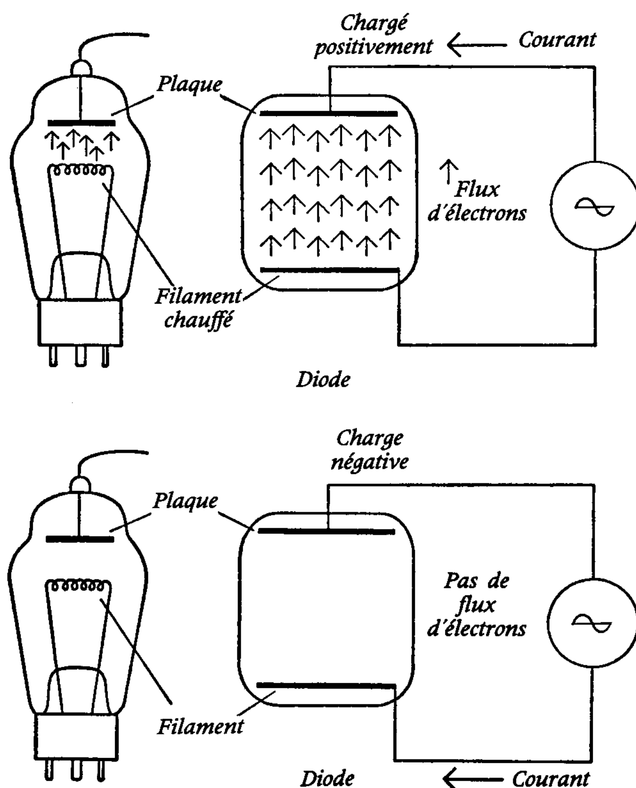
Les premiers fanatiques de radio devaient s'asseoir près de leur appareil en portant des écouteurs. On manquait de moyens de renforcer ou *amplifier* les signaux, et on trouva la réponse à ce problème dans une découverte d'Edison – la seule qu'il ait faite dans le domaine de la science « pure ».

Dans l'une de ses expériences ayant pour but d'améliorer l'ampoule électrique, Edison, en 1883, enferma un fil de métal dans une ampoule électrique au voisinage du filament. A sa grande surprise, de l'électricité circula du filament chauffé au fil de métal à travers l'intervalle de vide les séparant. Comme ce phénomène n'avait pas d'utilité pour ses recherches, Edison, homme pratique, le nota simplement dans ses cahiers et l'oublia. Mais l'*effet Edison* devint très important quand on eut découvert l'électron et qu'il fut devenu clair qu'un courant dans un espace vide de matière ne pouvait se concevoir que comme un flux d'électrons. Le physicien britannique Owen Williams Richardson montra, dans des

expériences réalisées entre 1900 et 1903, que les électrons se « vaporisaient » hors des filaments de métal chauffés dans le vide. Pour ce travail, il se vit décerner le prix Nobel en 1928.

En 1904, l'ingénieur électricien anglais John Ambrose Fleming trouva une brillante utilisation de l'effet Edison. Il entoura le filament d'une ampoule avec une pièce cylindrique de métal, appelée *plaque.* Cette plaque peut agir de deux manières. Si elle est chargée positivement, elle attire les électrons extraits du filament chauffé et crée ainsi un circuit où le courant électrique peut passer. Mais, si la plaque est chargée négativement, elle repousse les électrons et empêche donc toute circulation de courant. Supposons alors que la plaque soit reliée à une source de courant alternatif. Quand le courant passe dans un sens, la plaque se charge positivement et le courant passe dans le tube ; quand le courant alternatif change de sens, la plaque devient négativement chargée et aucun courant ne passe plus dans le tube. La plaque ne laisse donc passer le courant que dans une direction : elle convertit du courant alternatif en courant continu. Un tel tube agit comme une valve pour le courant ; aussi l'appelle-t-on simplement *valve,* bien que les scientifiques préfèrent le terme *diode,* parce qu'elle a deux électrodes – le filament et la plaque (figure 9.12).

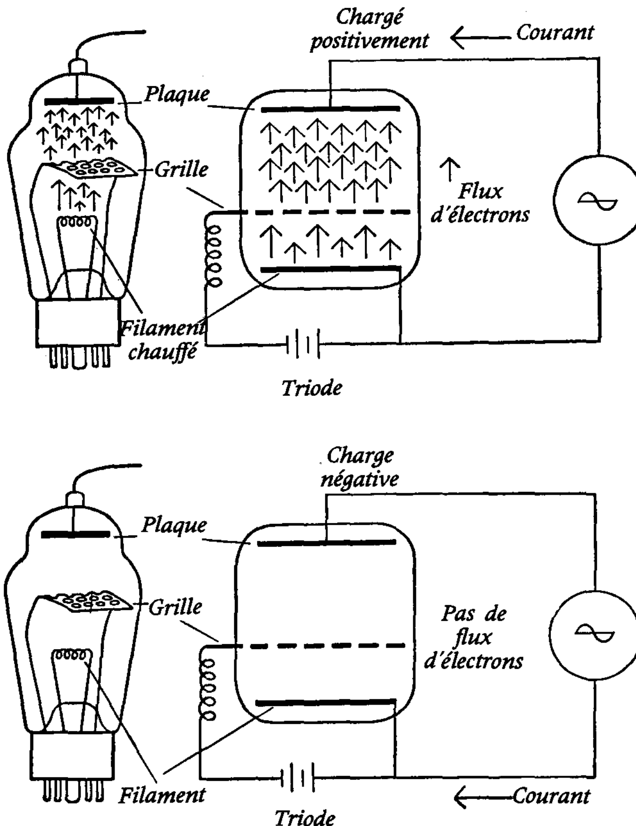
Figure 9.12. Principe de la diode à vide.



Le tube – ou *tube radio* puisque c'est dans ce domaine qu'il a été initialement utilisé – contrôle un courant d'électrons à travers le vide et non un courant électrique à travers un fil. Les électrons peuvent être contrôlés avec beaucoup plus de précision que le courant, de sorte que les tubes (et tous les dispositifs qui en découlèrent) constituèrent tout un nouveau domaine de *dispositifs électroniques* qui réalisaient des choses qu'aucun dispositif électrique simple ne pouvait réaliser. L'étude et l'utilisation des tubes et de leurs descendants ont reçu l'appellation d'*électronique.*

Le tube radio, sous sa forme la plus simple, sert de redresseur et remplace les cristaux utilisés jusqu'à cette époque, car il est beaucoup plus sûr. En 1907, l'inventeur américain Lee De Forest alla un peu plus loin. Il inséra une troisième électrode dans le tube, le transformant en *triode* (figure 9.13). La troisième électrode est une plaque perforée (la *grille*), placée entre le filament et la plaque. La grille attire les électrons et accélère leur écoulement du filament à la plaque (à travers les trous de la grille). Une petite augmentation de la charge positive de la grille entraîne une forte augmentation du flux d'électrons du filament à la plaque. Par conséquent, même la petite charge ajoutée par les faibles signaux radio augmente

Figure 9.13. Principe de la triode.



fortement le courant, et ce dernier reflète toutes les variations imposées par les ondes radio. En d'autres termes, la triode agit comme un amplificateur. Les triodes et d'autres modifications plus complexes du tube radio devinrent des pièces d'équipement essentielles, non seulement pour les récepteurs de radio, mais pour toutes sortes d'appareils électroniques.

Il fallait franchir une étape supplémentaire pour rendre complètement populaires les récepteurs de radio. Durant la Première Guerre mondiale, l'ingénieur électricien américain Edwin Howard Armstrong mit au point un dispositif pour abaisser la fréquence d'une onde radio. Il était destiné, à l'époque, à détecter les avions, mais après la guerre on l'utilisa dans les récepteurs de radio. Le *récepteur superhétérodyne* d'Armstrong permit de régler clairement un poste de radio sur une fréquence donnée en tournant un bouton, tandis qu'auparavant c'était une tâche fort compliquée que de régler le récepteur sur un large domaine de fréquences possibles. En 1921, une station de radio à Pittsburgh, aux États-Unis, commença à émettre des programmes réguliers. D'autres stations se montèrent très vite ; grâce aux contrôles du niveau sonore et de l'accord sur une station réduits à la manœuvre d'un bouton, les postes de radio devinrent extrêmement populaires. Vers 1927, on pouvait transmettre des conversations téléphoniques à travers les océans au moyen de la radio ; la *téléphonie sans fil* était née.

Il restait le problème des parasites. Les systèmes d'accord introduits par Marconi et ses successeurs minimisaient le bruit dû aux orages et autres sources d'électricité, mais ils ne l'éliminaient pas. C'est de nouveau Armstrong qui trouva la réponse au problème. Au lieu de la modulation d'amplitude, qui était sujette à des interférences avec les modulations d'amplitude aléatoires des sources de bruit, il introduisit en 1935 la *modulation de fréquence.* Il maintenait constante l'amplitude de l'onde radio porteuse et lui superposait une variation de sa fréquence. Quand le son avait une forte amplitude, l'onde porteuse voyait sa fréquence abaissée et vice versa. La modulation de fréquence (MF) élimina pratiquement les parasites et la radio MF se répandit beaucoup après la Deuxième Guerre mondiale, d'abord pour les programmes de musique classique.

LA TÉLÉVISION

La télévision a pris le relais de la radio, un peu comme les films parlants avaient pris le relais des films muets. Le précurseur technique de la télévision fut la transmission des images par fil, ce qui imposait de transformer une image en courant électrique. Un pinceau fin de lumière passait à travers l'image portée par un film photographique et atteignait une cellule photo-électrique. Aux endroits où le film était relativement opaque, un faible courant était engendré dans la cellule photo-électrique ; aux endroits où il était plus clair, un fort courant se formait dans la cellule. Le faisceau de lumière *balayait* lentement l'image de gauche à droite, ligne par ligne, et produisait ainsi un courant variable représentant toute l'image. On envoyait le courant dans des fils, et au lieu de destination il reproduisait l'image par un procédé inverse. Dès 1907, on transmit ainsi de tels *bélinogrammes* entre Londres et Paris.

La télévision transmet une image animée au lieu de simples photographies – que l'image animée provienne d'une prise en « direct » ou d'un

film. La transmission doit être extrêmement rapide, puisque l'action doit être balayée très rapidement. La structure lumineuse de l'image est convertie en une structure d'impulsions électriques au moyen d'une caméra utilisant, à la place de la pellicule, un revêtement de métal qui émet des électrons quand la lumière l'atteint.

Une espèce de télévision fut mise en démonstration en 1926 pour la première fois par l'inventeur écossais John Logie Baird. Cependant, la première caméra de télévision d'usage pratique fut l'*iconoscope*, breveté en 1938 par l'inventeur américain né en Russie, Vladimir Kosma Zworykin. Dans l'iconoscope, l'arrière de la caméra est revêtu d'un grand nombre de petites gouttelettes de césium-argent. Chacune émet des électrons dès que le faisceau lumineux la balaye, et cela proportionnellement à l'intensité de la lumière. L'iconoscope fut remplacé plus tard par l'*image orthicon* – un perfectionnement dans lequel l'écran en césium-argent est assez fin pour que les électrons émis puissent le traverser et atteindre une fine plaque de verre qui émet plus d'électrons. Cette amplification augmente la sensibilité de la caméra à la lumière, de sorte qu'on n'a plus besoin d'un très fort éclairage.

Le récepteur de télévision utilise une sorte de tube à rayons cathodiques. Un faisceau d'électrons émis par un filament (le *canon à électrons*) heurte un écran revêtu d'une substance fluorescente, qui émet de la lumière dont l'intensité est proportionnelle à celle du faisceau d'électrons. Des paires d'électrodes contrôlant la direction du faisceau le font balayer l'écran de gauche à droite, en une série de plusieurs centaines de lignes horizontales, chacune d'elles étant légèrement au-dessous de la précédente. De cette façon, le « balayage » complet d'une image sur l'écran prend 1/30 de seconde aux États-Unis, 1/25 de seconde en Europe. Le faisceau balaye donc des images successives à la cadence de vingt-cinq ou trente par seconde. A aucun moment il n'y a sur l'écran plus d'un point allumé (brillant ou sombre selon le cas) ; pourtant, grâce à la persistance de la vision, nous voyons non seulement des images complètes, mais une suite ininterrompue de mouvements.

Dans les années vingt, il y eut des émissions expérimentales de télévision, mais la télévision ne fut commercialisée qu'en 1947. Depuis, elle domine pour ainsi dire le secteur des loisirs.

Dans le milieu des années cinquante, on lui a ajouté deux perfectionnements. En utilisant trois types de matériaux fluorescents sur l'écran, chacun élaboré pour réagir au faisceau en rouge, bleu ou vert, on est parvenu à mettre au point la télévision en couleurs. Et le *magnétoscope*, type d'enregistreur ayant certaines similitudes avec la piste sonore d'un film, a rendu possible la reproduction des émissions avec une meilleure qualité que celle obtenue avec un film de cinéma.

LE TRANSISTOR

Dans les années quatre-vingts, le monde est arrivé à l'*âge de la cassette*. De même que de petites cassettes peuvent dérouler et réenrouler leur ruban pour reproduire la musique en haute fidélité – branchées sur des piles, si nécessaire, afin que l'on puisse se promener ou faire le ménage avec des écouteurs sur la tête, chacun écoutant sa propre musique –, il existe des *vidéocassettes*, qui peuvent reproduire des films de n'importe quel type

sur un poste de télévision ou enregistrer des programmes au moment de l'émission pour les projeter ensuite.

Le tube à vide, le cœur de tous les dispositifs électroniques, était devenu depuis longtemps un facteur limitant les performances. Habituellement, les composants d'un dispositif sont constamment améliorés en rendement au cours du temps, c'est-à-dire qu'on augmente leur puissance et leur facilité d'emploi et qu'on réduit leur taille et leur masse (procédé parfois appelé *miniaturisation*). Mais le tube à vide devenait un obstacle à la course à la miniaturisation, parce qu'il devait rester suffisamment gros pour offrir au vide un volume tel que les divers composants qu'il contenait ne subissent pas de fuites électriques pour cause d'espacement trop petit.

Il présentait en outre d'autres inconvénients. Le tube pouvait se casser ou perdre son vide et, dans les deux cas, devenir inutilisable. (On remplaçait fréquemment les tubes dans les premiers postes de radio et de télévision et, dans ce dernier cas particulièrement, il aurait presque fallu un réparateur à demeure.) En outre, les tubes ne pouvaient fonctionner que si les filaments étaient suffisamment chauffés ; il leur fallait donc un fort courant et il y avait un temps d'attente pour que le poste « chauffe ». Et puis, presque par accident, surgit une solution inattendue. Dans les années quarante, plusieurs chercheurs des laboratoires Bell, aux États-Unis, s'intéressaient de plus en plus à des substances connues sous le nom de *semi-conducteurs.* Ces éléments, comme le silicium et le germanium, conduisent moyennement l'électricité, et le problème se posait de comprendre pourquoi. Les chercheurs de Bell découvrirent que la conductivité que présentent ces substances était augmentée par des traces d'impuretés mélangées à l'élément en question.

Considérons un cristal de germanium. Chaque atome possède quatre électrons sur sa couche extérieure ; et dans l'arrangement régulier des atomes du cristal, chacun des quatre électrons s'apparie avec un électron d'un atome voisin de germanium, de sorte que tous les électrons sont appariés en liaisons stables. Comme cet arrangement est similaire à celui du diamant, le germanium, le silicium et d'autres substances analogues sont appelées *adamantines,* d'un mot ancien pour « diamant ». Si l'on introduit un peu d'arsenic dans cet arrangement adamantin, le schéma devient plus compliqué. L'arsenic possède cinq électrons sur sa dernière couche. Un atome d'arsenic, prenant la place d'un atome de germanium dans le cristal, aura la possibilité d'apparier quatre de ses cinq électrons avec les atomes voisins, mais le cinquième ne trouvera pas d'électron auquel s'apparier : il sera donc peu lié. Maintenant, si l'on applique une tension électrique à ce cristal, l'électron peu lié va se déplacer dans la direction de l'électrode positive. Il ne va pas se déplacer aussi librement que le ferait un électron dans un métal, mais le cristal conduira l'électricité mieux qu'un isolant comme le soufre ou le verre. Ceci n'est pas très étonnant, mais venons-en à un cas quelque peu plus étrange. Ajoutons un peu de bore, au lieu de l'arsenic, au germanium. L'atome de bore n'a que trois électrons sur sa couche extérieure. Ces trois électrons peuvent s'apparier avec les électrons de trois atomes voisins de germanium. Mais qu'arrive-t-il à l'électron du quatrième atome de germanium voisin de l'atome de bore ? Cet électron est apparié avec un trou ! Le mot *trou* est utilisé à dessein, car le site où l'électron trouverait un partenaire dans un cristal de germanium se comporte en fait comme une lacune. Si on applique une tension au cristal contaminé par du bore, le plus proche

électron voisin, attiré vers l'électrode positive, ira dans ce trou. Ce faisant, il laisse un trou où il était, et l'électron suivant le plus éloigné de l'électrode positive va dans ce nouveau trou. De sorte que le trou va à coup sûr se diriger vers l'électrode négative, en se déplaçant exactement comme un électron, mais dans la direction opposée. En résumé, il est devenu porteur de courant électrique.

Pour que tout fonctionne bien, il faut que le cristal soit presque parfaitement pur avec juste la quantité voulue d'impureté spécifiée (c'est-à-dire d'arsenic ou de bore). On dit que le semi-conducteur germanium-arsenic avec un électron baladeur est du *type n* (pour « négatif »). Le semi-conducteur germanium-bore, avec un trou baladeur, qui agit comme s'il était positivement chargé, est dit du *type p* (pour « positif »).

A la différence des conducteurs ordinaires, la résistance électrique des semi-conducteurs baisse quand la température augmente, parce que des températures plus élevées affaiblissent l'emprise des atomes sur les électrons, permettant à ces derniers de dériver plus librement. (Dans les conducteurs métalliques, les électrons sont déjà assez libres à température ordinaire. Augmenter la température introduit plus de mouvement aléatoire et gêne leur déplacement en réponse à un champ électrique.) En déterminant la résistance d'un semi-conducteur, on peut mesurer des températures qui sont trop élevées pour être aisément mesurées par d'autres méthodes. On appelle *thermistances* de tels semi-conducteurs servant à mesurer les températures.

Mais les semi-conducteurs peuvent faire bien plus si on les associe astucieusement. Supposons que nous fabriquions un cristal de germanium avec une moitié de type p et l'autre moitié de type n. Si nous relions le côté n à une électrode négative et le côté p à une électrode positive, les électrons du côté de type n se déplaceront à travers le cristal vers l'électrode positive, tandis que les trous du côté de type p iront en direction opposée vers l'électrode négative. Un courant circule donc à travers le cristal. Inversons maintenant la situation – c'est-à-dire que nous connectons le côté de type n à l'électrode positive, et celui du type p à l'électrode négative. Cette fois-ci, les électrons du côté n vont vers l'électrode positive – c'est-à-dire qu'ils s'éloignent du côté p – et les trous du côté p s'éloignent de façon similaire du côté n. Il en résulte que les régions frontières à la jonction entre les côtés n et p perdent leurs électrons et trous libres, ce qui crée une coupure dans le circuit, et aucun courant ne passe.

En résumé, nous avons maintenant un dispositif qui peut jouer le rôle d'un redresseur. Si nous branchons ce double cristal sur du courant alternatif, le cristal laissera passer le courant dans un sens et pas dans l'autre. Donc, du courant alternatif sera converti en courant continu. Le cristal sert de diode, exactement comme un tube à vide (ou *valve*).

En un sens, l'électronique avait bouclé la boucle. Le tube avait remplacé le cristal et maintenant, le cristal avait remplacé le tube – mais c'était une nouvelle sorte de cristal, bien plus délicate et sûre que celle que Braun avait introduite près d'un demi-siècle auparavant.

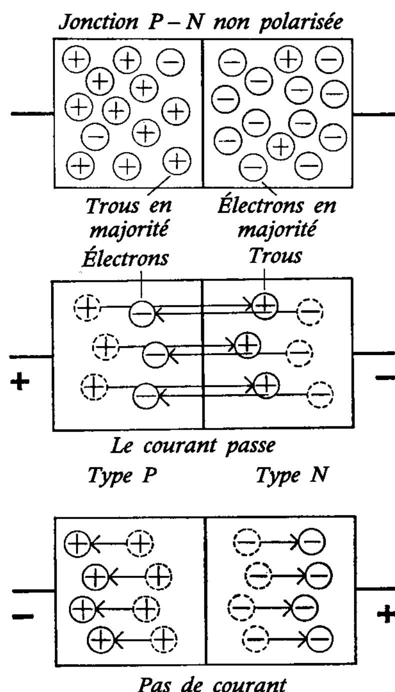
Le nouveau cristal avait des avantages impressionnants sur le tube. Il ne nécessitait pas de vide, il pouvait donc être petit. Il ne se cassait pas, n'était pas sujet aux fuites. Comme il fonctionnait à la température ambiante, il demandait très peu de courant et aucun temps de chauffage. Il n'avait que des avantages et aucun inconvénient, pourvu seulement qu'on pût le fabriquer de façon assez économique et précise.

Puisque les nouveaux cristaux étaient solides, ils ouvraient la voie à ce qu'on allait appeler *l'électronique à l'état solide.* Le nouveau dispositif fut appelé *transistor* (selon l'idée de John Robinson Pierce, des Laboratoires Bell), parce qu'il *trans*fère un signal à travers une résistance (en anglais re*sistor*) (figure 9.14).

En 1948, toujours aux Laboratoires Bell, William Bradford Shockley, Walter Houser Brattain et John Bardeen réussirent à produire un transistor qui pouvait servir d'amplificateur. C'était un cristal de germanium avec une fine section de type p mise en sandwich entre deux bouts de type n. C'était en fait une triode avec l'équivalent d'une grille entre le filament et la plaque. Par contrôle de la charge positive dans le centre de type p, on pouvait envoyer des trous à travers les jonctions de façon à contrôler le flux d'électrons. En outre, une faible variation du courant dans le centre de type p causait une forte variation du courant à travers ce système semi-conducteur. La triode à semi-conducteur pouvait donc servir d'amplificateur exactement comme le faisait un tube triode à vide. Shockley, Brattain et Bardeen reçurent le prix Nobel de physique en 1956.

Même si les semi-conducteurs marchaient bien en théorie, leur usage pratique nécessitait certains progrès techniques supplémentaires – comme c'est toujours le cas en sciences appliquées. Le rendement des transistors dépend beaucoup de l'usage de matériaux de pureté extrêmement élevée, afin que l'on puisse soigneusement contrôler la nature et la concentration des impuretés que l'on ajoute.

Figure 9.14. Principe du transistor à jonction.



Heureusement, William Gardner Pfann introduisit en 1952 la technique du *raffinage par fusion de zone.* Un barreau de germanium, par exemple, est placé à l'intérieur d'un élément de chauffage circulaire, qui commence à fondre une section du barreau. On tire celui-ci à travers l'élément de chauffage afin que la zone fondue se déplace le long du barreau. Les impuretés du barreau tendent à rester dans la zone fondue et sont littéralement balayées jusqu'aux extrémités du barreau. Après quelques passes de ce genre, la partie principale du barreau de germanium est d'une pureté sans précédent.

Vers 1953, on se mit à utiliser de minuscules transistors dans des appareils pour sourds, ce qui permit de rendre ces derniers assez petits pour pouvoir être installés dans l'oreille. Le transistor continuant à se développer de façon constante, il put très vite fonctionner à de plus hautes fréquences, supporter des températures de plus en plus élevées et voir sa taille de plus en plus réduite. Il est finalement devenu si minuscule qu'on a cessé d'utiliser des transistors individuels. On a usiné à l'échelle microscopique, par attaque chimique, de petites « puces » de silicium pour obtenir des *circuits intégrés* qui remplissaient des fonctions pour lesquelles il aurait fallu un très grand nombre de tubes. Dans les années soixante-dix, les puces étaient devenues encore plus petites, et on les a appelées *micropuces.*

De tels dispositifs minuscules à l'état solide, maintenant d'un usage universel, représentent peut-être la plus étonnante de toutes les révolutions scientifiques de l'histoire de l'humanité. Ils ont rendu possible la miniaturisation des radios ; ils ont permis d'entasser dans les satellites et les sondes spatiales un matériel aux énormes potentialités ; surtout, ils ont permis le développement d'ordinateurs de plus en plus petits, de plus en plus économiques et de plus en plus polyvalents, et même de robots, dans les années quatre-vingts. Ces deux derniers sujets seront traités plus loin, au chapitre 17.

Les masers et les lasers

LES MASERS

Une autre avance récente et gigantesque a commencé avec des recherches concernant la molécule d'ammoniac (NH_3). On peut se la représenter avec les trois atomes d'hydrogène occupant les sommets d'un triangle équilatéral, tandis que l'atome d'azote est à quelque distance au-dessus du centre de symétrie du triangle. La molécule d'ammoniac peut vibrer : l'atome d'azote se déplace alors en traversant le plan du triangle vers une position symétrique de l'autre côté, puis revient à sa position initiale, et ainsi de suite. La molécule d'ammoniac peut vibrer ainsi avec une fréquence propre de 24 milliards de cycles par seconde.

Cette période de vibration est extrêmement constante, beaucoup plus que celle de n'importe quel dispositif artificiel, plus constante même que les périodes des mouvements des corps célestes. Ces molécules en vibration peuvent être utilisées pour asservir des courants électriques qui commanderont à leur tour des dispositifs à mesurer le temps avec une précision sans

précédent – la première démonstration en a été faite en 1949 par le physicien américain Harold Lyons. Vers 1955, de telles *horloges atomiques* surpassaient de fort loin tous les chronomètres ordinaires. On a atteint des précisions sur la mesure du temps de une seconde pour 1 700 000 années en utilisant des atomes d'hydrogène.

La molécule d'ammoniac, au cours de ces vibrations, émet un faisceau de rayonnement électromagnétique ayant une fréquence de 24 milliards de cycles par seconde. Ce rayonnement a une longueur d'onde de 1,25 centimètre et appartient au domaine des hyperfréquences. Une autre façon de considérer la question est d'imaginer la molécule d'ammoniac comme pouvant occuper l'un ou l'autre de deux niveaux d'énergie ayant une différence d'énergie égale à celle d'un photon de 1,25 centimètre de longueur d'onde. Si la molécule d'ammoniac tombe du niveau d'énergie le plus élevé vers le plus bas, elle émet un photon possédant exactement cette énergie. Si une molécule dans le niveau le plus bas d'énergie absorbe un photon de cette énergie, elle monte dans le niveau d'énergie le plus élevé.

Mais que va-t-il se passer si une molécule d'ammoniac déjà dans le niveau d'énergie le plus élevé est exposée à de tels photons ? Dès 1917, Einstein avait prédit que, si un photon ayant juste l'énergie requise heurtait une molécule du niveau supérieur, celle-ci serait repoussée vers le niveau le plus bas en émettant un photon ayant exactement la même énergie et la même direction de mouvement que celles du photon incident. Il y aurait donc deux photons identiques là où il n'y en avait qu'un auparavant. Cette théorie a reçu sa confirmation expérimentale en 1924.

L'ammoniac exposé aux hyperfréquences pouvait donc subir deux changements : des molécules pouvaient être pompées du niveau d'énergie le plus bas vers le niveau le plus élevé, ou induites à tomber du plus haut niveau vers le plus bas. Dans les conditions normales, c'est le premier processus qui devait être prédominant car seul un très faible pourcentage des molécules d'ammoniac se trouverait, à tout instant, dans le niveau d'énergie le plus élevé.

Supposons, cependant, que l'on ait trouvé une méthode pour placer toutes ou presque toutes les molécules dans le niveau d'énergie supérieur. C'est alors le mouvement du niveau le plus haut vers le plus bas qui prédominerait. D'ailleurs, il se passerait quelque chose de très intéresant. Le faisceau incident de rayonnement hyperfréquences fournirait un photon qui induirait le mouvement d'une molécule vers le niveau bas. Un second photon serait donc émis et il se déplacerait en compagnie du précédent jusqu'à heurter deux molécules, d'où l'émission de deux nouveaux photons. Ces quatre photons entraîneraient l'émission de quatre autres photons et ainsi de suite. Le photon initial pourrait donc déclencher une véritable avalanche de photons, ayant tous exactement la même énergie et allant exactement dans la même direction.

En 1953, le physicien américain Charles Hard Townes conçut une méthode pour isoler des molécules d'ammoniac dans le niveau d'énergie le plus haut et les soumettre à une excitation par des photons hyperfréquences d'énergie adaptée. Avec peu de photons à l'entrée du système, on en a recueilli une foule à la sortie. Le rayonnement incident avait donc été fortement amplifié.

Ce processus fut décrit comme une « amplification hyperfréquences par une émission stimulée de rayonnement », en anglais « *m*icrowave

*a*mplification by *s*timulated *e*mission of *r*adiation ». On a donc appelé l'instrument un *maser.*

On conçut bientôt des masers à état solide – avec des solides où les électrons pouvaient avoir deux niveaux d'énergie à occuper. Les premiers masers, tant à gaz que solides, ne fonctionnaient que par intermittence, c'est-à-dire qu'il fallait d'abord les pomper dans l'état d'énergie le plus élevé, puis les stimuler. Après une courte impulsion d'émission de rayonnement, il ne se passait plus rien jusqu'à ce que le processus de pompage se soit répété.

Pour contourner cet obstacle, il vint à l'esprit du physicien américain d'origine hollandaise Nicolaas Bloembergen d'utiliser un système à trois niveaux. Si le matériau choisi pour le noyau du maser pouvait avoir des électrons dans l'un des trois niveaux – un niveau inférieur, un central et un supérieur –, alors le pompage et l'émission pouvaient avoir lieu simultanément. Les électrons sont pompés du niveau le plus bas vers le plus élevé. Une fois dans le niveau le plus élevé, une stimulation adéquate peut les faire tomber, d'abord vers le niveau du milieu, puis vers le niveau inférieur. Comme des photons de différentes énergies sont nécessaires pour le pompage et l'émission stimulée, les deux processus n'interfèrent pas l'un avec l'autre. On peut donc finalement avoir un maser continu.

En tant qu'amplificateurs hyperfréquences, les masers peuvent être utilisés comme détecteurs très sensibles en radioastronomie, où des signaux hyperfréquences extrêmement faibles venant de l'espace doivent être fortement intensifiés avec une grande fidélité par rapport aux caractéristiques du rayonnement d'origine. (Reproduire sans perte des caractéristiques originales, c'est reproduire avec peu de « bruit ». Les masers sont extrêmement « peu bruyants » dans ce sens du mot.) Ils ont aussi prouvé leur utilité dans l'espace. A bord du satellite soviétique *Cosmos 97,* lancé le 30 novembre 1965, se trouvait un maser qui fonctionna parfaitement.

Townes partagea en 1964 le prix Nobel de physique avec deux physiciens soviétiques, Nicolai Gennedievitch Basov et Alexandre Mikhailovitch Prochorov, qui avaient indépendamment travaillé à la théorie du maser.

LES LASERS

En principe, la technique du maser pouvait être appliquée aux ondes électromagnétiques de longueur d'onde quelconque, et donc à celles de la lumière visible. Townes dirigea l'effort de recherche vers les longueurs d'onde de la lumière visible en 1958. Un tel maser peut s'appeler *maser optique.* Mais le processus est en fait une « amplification de lumière par émission stimulée de rayonnement », en anglais « *l*ight *a*mplification by *s*timulated *e*mission of *r*adition », d'où le terme populaire *laser* (figure 9.15).

Le premier laser à avoir fonctionné fut celui construit en 1960 par le physicien américain Theodore Harold Maiman. Maiman utilisa un barreau de rubis synthétique, c'est-à-dire constitué essentiellement d'alumine (oxyde d'aluminium), avec une trace d'oxyde de chrome. Si on expose le barreau de rubis à la lumière, les électrons des atomes de chrome sont pompés vers des niveaux plus élevés et après un court laps de temps, ils commencent à retomber. Les rares premiers photons de lumière émise (avec une longueur d'onde de 694,3 millimicrons) stimulent la production d'autres photons identiques, et le barreau émet un faisceau de lumière

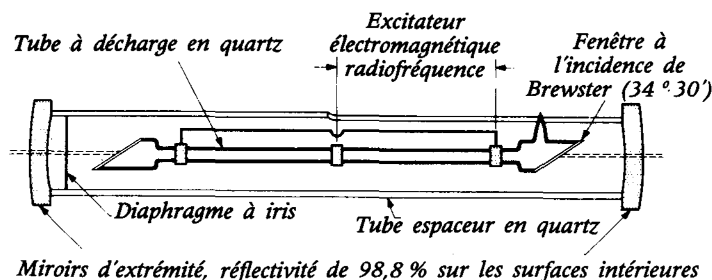


Figure 9.15. Laser à fonctionnement continu avec miroirs concaves et fenêtres à l'incidence de Brewster sur le tube à décharge. Le tube est rempli d'un gaz dont les atomes sont pompés vers des états de haute énergie par excitation électromagnétique. Ces atomes sont alors stimulés par l'introduction d'un faisceau de lumière, et ils émettent de l'énergie à une longueur d'onde donnée. Agissant à la manière d'un tuyau d'orgue, la cavité résonnante voit s'établir un train d'ondes cohérentes entre ses miroirs. Le faisceau fin qui s'échappe est le rayon laser. D'après un dessin paru aux États-Unis dans la revue *Science* du 9 octobre 1964.

rouge quatre fois plus intense que la lumière à la surface du Soleil. Avant la fin de l'année 1960, des lasers continus ont été montés par un physicien iranien, Ali Javan, qui travaillait aux Laboratoires Bell. Il utilisait comme milieu actif un mélange gazeux (hélium et néon).

Le laser fournit une lumière sous une forme complètement nouvelle, plus intense qu'aucune autre lumière jamais produite, et également plus fortement monochromatique (une seule longueur d'onde). Mais il y a plus.

La lumière ordinaire, produite par toute autre méthode – du feu de bois au Soleil en passant par la luciole – consiste en paquets d'ondes relativement courts. On peut les représenter comme des petits morceaux d'ondes pointant dans toutes les directions. La lumière ordinaire est faite d'un nombre infini de ceux-ci.

La lumière produite par un laser, au contraire, consiste en photons ayant tous la même énergie et se déplaçant tous dans la même direction. Les paquets d'ondes ont donc tous la même fréquence ; et, comme ils sont alignés côte à côte très précisément, ils se fondent, pour ainsi dire, ensemble. La lumière apparaît comme constituée de longs paquets d'ondes d'amplitude (hauteur) et de fréquence (largeur) constantes. C'est de la *lumière cohérente :* les paquets d'ondes semblent se coller les uns contre les autres. Les physiciens savaient déjà créer des rayonnements cohérents pour de grandes longueurs d'onde. Pour la lumière, cependant, il fallut attendre jusqu'en 1960.

Le laser a été conçu, en outre, pour que la tendance naturelle des photons à aller dans la même direction soit accentuée. Les deux extrémités du barreau de rubis étaient usinées très précisément et argentées pour servir de miroirs plans. Les photons émis allaient et venaient le long du barreau, heurtant plus de photons à chaque passage, jusqu'à atteindre une intensité suffisante pour sortir par l'extrémité qui était la plus légèrement argentée. Ceux qui sortaient par là étaient précisément ceux qui avaient été émis dans une direction exactement parallèle à celle du long axe du barreau,

car ils avaient fait de nombreux va-et-vient, heurtant les extrémités réfléchissantes de façon répétée. Si un photon ayant la bonne énergie arrivait à entrer dans le barreau dans une direction différente (même une direction très légèrement différente) et déclenchait un train de photons stimulés dans cette direction différente de l'axe, ceux-ci sortiraient vite par les côtés du barreau après quelques réflexions au plus.

Un faisceau de lumière laser est constitué d'ondes cohérentes si parallèles qu'il peut parcourir de longues distances sans trop diverger. Il peut être focalisé assez finement pour chauffer une cafetière à un millier de kilomètres. Des faisceaux laser ont même atteint la Lune, en 1962, s'étalant sur un diamètre de trois kilomètres après avoir parcouru plus de 300 000 kilomètres !

Une fois le laser mis au point, on songea vite à lui apporter de multiples perfectionnements. En quelques années, on développa des lasers individuels capables de produire de la lumière cohérente à une centaine de longueurs d'onde différentes, du proche ultraviolet à l'infrarouge lointain. L'effet laser est obtenu à partir d'une large variété de solides (oxydes métalliques, fluorures, tungstates, semi-conducteurs), de liquides et de gaz. Chaque variété a ses avantages et ses inconvénients.

En 1964, le physicien américain Jerome V.V. Kasper mit au point le premier *laser chimique.* Dans un tel laser, la source d'énergie est une réaction chimique (dans le cas du premier de ces lasers, la dissociation de CF_3I par une impulsion de lumière). L'avantage du laser chimique sur les autres est que la réaction chimique fournissant l'énergie peut être incorporée dans le laser lui-même, et qu'aucune source extérieure d'énergie n'est nécessaire. C'est comme un dispositif alimenté par pile comparé à un autre qui doit être branché à une prise électrique. Il y a là un grand avantage dû au fait que l'appareil est portatif, sans compter que les lasers chimiques semblent avoir un rendement plus élevé que les autres types de lasers (12 % ou plus, comparé à 2 % ou moins).

Des *lasers organiques,* dans lesquels un colorant organique complexe est utilisé comme source de lumière cohérente, ont été mis au point pour la première fois en 1966 par John R. Lankard et Peter Sorokin. La complexité de la molécule rend possible la production de lumière par une grande variété de réactions électroniques et donc avec une grande variété de longueurs d'onde. On peut *accorder* un seul laser organique pour obtenir l'une des longueurs d'onde dans un certain domaine plutôt que de se trouver limité à une longueur d'onde unique, comme c'est le cas avec les autres lasers.

La finesse du faisceau de lumière laser permet de focaliser une grande quantité d'énergie sur une surface extrêmement réduite ; sur cette surface, la température atteint des niveaux très élevés. Le laser peut vaporiser des métaux pour les besoins de la recherche et l'analyse spectrale rapide. Il peut souder et découper des substances à haut point de fusion et y percer des trous d'une forme bien déterminée. En dirigeant dans l'œil des faisceaux lasers, des chirurgiens ont réussi à réparer des décollements de rétine, en opérant assez vite pour que les tissus environnants n'aient pas le temps d'être affectés par la chaleur. De façon analogue, on a utilisé des lasers pour détruire des tumeurs.

Pour démontrer le vaste champ d'application du laser, Arthur L. Shawlow a mis au point la très simple (mais impressionnante) *gomme à laser,* qui en un flash très bref et intense vaporise l'encre de caractères

tapés à la machine sur du papier sans pour autant le noircir ; à l'autre extrême, les *interféromètres à laser* permettent des mesures d'une précision sans précédent. Les contraintes telluriques à la surface du globe terrestre peuvent être détectées rien qu'en observant la manière dont sont modifiées les franges d'interférences produites avec des lasers. Des mouvements telluriques peuvent être ainsi détectés avec une précision d'une partie pour mille milliards. Les premiers hommes à être allés sur la Lune y ont laissé un réflecteur laser conçu pour renvoyer vers la Terre des faisceaux lasers. Avec une telle méthode, la distance Terre-Lune a pu être déterminée avec une plus grande précision que celle d'un point à un autre sur la Terre.

Une des applications, qui jouit d'un grand engouement depuis le début, a été l'utilisation des faisceaux lasers comme faisceaux porteurs dans les communications. La haute fréquence de la lumière cohérente, comparée à celle des ondes radio cohérentes utilisées aujourd'hui pour la radio et la télévision, porte en germe l'espoir de faire tenir plusieurs milliers de canaux de transmission dans l'espace occupé actuellement par un seul. On pourrait même imaginer que chaque être humain sur la Terre puisse avoir sa longueur d'onde personnelle. Il faut pour cela que la lumière laser soit modulée. Les courants électriques variables produits à partir du son doivent être traduits en variations de lumière laser (variation d'amplitude ou variation de fréquence), qui peuvent à leur tour être utilisées pour produire ailleurs des courants électriques variables. On développe actuellement de tels systèmes.

Comme la lumière est plus sensible que les ondes radio aux interférences dues aux nuages, à la brume, au brouillard et à la poussière, il sera nécessaire de guider la lumière laser à travers des conduits contenant des lentilles (pour reconcentrer le faisceau à intervalles réguliers) et des miroirs (pour le réfléchir lors des coudes). On a cependant mis au point des *lasers à dioxyde de carbone* (CO_2), qui produisent des faisceaux lasers continus d'une puissance sans précédent et qui sont situés assez loin dans l'infrarouge pour être peu affectés par l'atmosphère. La communication atmosphérique semble donc être possible. Mais il semble beaucoup plus pratique dans l'immédiat d'envoyer des faisceaux lasers modulés dans des *fibres optiques,* petits fils de verre très transparents plus fins qu'un cheveu humain, pour remplacer les fils de cuivre isolés dans les communications téléphoniques. Le verre est infiniment moins cher et plus répandu que le cuivre, et il peut transporter bien plus d'informations au moyen de la lumière laser. Déjà, en de nombreux endroits, les câbles de fil de cuivre encombrants laissent la place aux faisceaux de fibres optiques, qui le sont beaucoup moins.

Une application encore plus fascinante des faisceaux lasers, qui se répand beaucoup, concerne une nouvelle sorte de photographie. Dans la photographie ordinaire, un faisceau de lumière ordinaire réfléchi par un objet tombe sur un film photographique. Ce qui est enregistré, c'est la section transverse de la lumière, c'est-à-dire en aucune façon toute l'information qu'elle peut contenir.

Supposons qu'un faisceau de lumière soit séparé en deux parties. L'une d'elles heurte un objet et est réfléchie avec toutes les irrégularités que cet objet lui impose. La seconde partie est réfléchie sur un miroir sans irrégularités. Les deux parties se rencontrent sur le film photographique, qui enregistre l'interférence des diverses longueurs d'onde. Théoriquement, l'enregistrement de cette interférence devrait contenir toutes les données

concernant chaque faisceau de lumière. La photographie qui a enregistré cette figure d'interférences semble blanche quand on la développe ; mais si on envoie de la lumière sur le film, si elle le traverse et explore donc les caractéristiques de l'interférence, cela produit une image contenant l'information complète. Cette image est tridimensionnelle comme l'était la surface de l'objet par lequel la lumière avait été réfléchie, et on peut d'ailleurs en prendre des photographies ordinaires sous divers angles : elles montreront des changements dans la perspective.

Cette idée avait été avancée pour la première fois par le physicien britannique d'origine hongroise Dennis Gabor en 1947, alors qu'il essayait de trouver des méthodes pour affiner les images produites par des microscopes électroniques. Il appela cette méthode *holographie*, d'un mot latin signifiant « écriture tout entière de la main de l'auteur ».

Bien que l'idée de Gabor soit, du point de vue théorique, parfaitement correcte, on ne peut la mettre en œuvre avec de la lumière ordinaire. En effet, avec des longueurs d'onde de toutes les valeurs se déplaçant dans toutes les directions, les franges d'interférences produites par les deux faisceaux de lumière sont si désordonnées qu'on ne recueille aucune information. On aboutit à un million d'images floues, toutes superposées et dans des positions légèrement différentes.

L'introduction de la lumière laser a changé tout cela. En 1965, Emmet N. Leith et Juris Upatnieks, de l'université de Michigan, aux États-Unis, ont produit les premiers hologrammes. Depuis lors la technique a tant progressé que l'holographie en couleurs est devenue possible, et que les franges d'interférences photographiées peuvent être observées en lumière ordinaire. La *microholographie* promet d'ajouter une nouvelle dimension (littéralement) aux recherches biologiques ; nul ne peut dire jusqu'où cela ira.

Chapitre 10

Le réacteur nucléaire

L'énergie

Les progrès rapides de la technique au XXe siècle ont eu pour contrecoup une gigantesque augmentation de notre consommation en ressources énergétiques venant de la Terre. Si les nations sous-développées, avec leurs milliards d'habitants, rejoignent les contrées industrialisées à haut niveau de vie, le taux de consommation de carburant augmentera de façon encore plus spectaculaire. Où trouverons-nous les quantités d'énergie nécessaires à la survie de notre civilisation ?

Nous avons déjà vu la plupart des ressources en bois disparaître de la Terre. Le bois était notre premier combustible. Au début de l'ère chrétienne, une partie importante de la Grèce, de l'Afrique du Nord et du Proche-Orient fut systématiquement déboisée, d'une part pour le combustible, d'autre part pour permettre le développement de l'élevage et de l'agriculture. La mise en coupe réglée des forêts fut un double désastre. Non seulement elle détruisit les réserves de bois, mais la terrible mise à nu de la campagne entraîna un anéantissement plus ou moins définitif de la fertilité. La plupart de ces régions, autrefois berceaux de cultures avancées, sont aujourd'hui stériles et improductives, peuplées par une population appauvrie et en mauvaise santé.

Le Moyen Age a vu le déboisement progressif de l'Europe de l'Ouest, et les temps modernes le déboisement, beaucoup plus rapide, du continent nord-américain. Il ne subsiste presque aucun grand espace de forêt vierge dans les zones tempérées du monde, sauf au Canada et en Sibérie.

LE CHARBON ET LE PÉTROLE, COMBUSTIBLES FOSSILES

Le charbon et le pétrole ont pris la place du bois comme combustibles. On trouve mention du charbon chez le botaniste grec Théophraste en 200 av. J.-C., mais les premières traces d'exploitation minière du charbon en Europe ne datent que du XIIe siècle. Vers le XVIIe siècle, l'Angleterre, alors fort déboisée et à court de bois pour sa flotte, commença à se convertir à l'usage à grande échelle du charbon comme combustible, inspirée peut-être par le fait que les Hollandais avaient commencé à creuser pour chercher du charbon. (Ils n'étaient pas les premiers : Marco Polo, dans le célèbre récit de ses voyages en Chine, en 1298, avait décrit l'usage du charbon comme combustible dans ce pays, qui était alors le plus avancé du monde techniquement parlant.)

Vers 1660, l'Angleterre produisait deux millions de tonnes de charbon chaque année, soit plus de 80 % de tout le charbon alors produit dans le monde.

Au début il était principalement utilisé comme combustible domestique ; mais en 1603 un Anglais, Hugh Platt, découvrit que si le charbon était chauffé en l'absence d'oxygène, les matériaux goudronneux et bitumineux qu'il contenait se dégageaient et brûlaient. Il restait alors du carbone presque pur, et on appela *coke* ce résidu.

Au début le coke n'était pas de haute qualité. On l'améliora peu à peu et finalement on put l'utiliser à la place du charbon de bois pour traiter le minerai de fer. Le coke brûlait à haute température, ses atomes de carbone se combinaient avec les atomes d'oxygène du minerai de fer, et il restait le fer métallique. En 1709, un Anglais, Abraham Darby, commença à utiliser le coke à grande échelle pour la fabrication du fer. Quand la machine à vapeur arriva, le charbon servit à chauffer l'eau pour la faire bouillir ; c'est ainsi que la Révolution industrielle fut mise sur les rails.

La tendance était plus lente ailleurs. Même en 1800, le bois fournissait 94 % des besoins en combustible aux États-Unis, jeunes encore et riches en forêts. En 1885, cependant, le bois fournissait 50 % seulement des besoins en combustible et vers 1980, moins de 3 %. La balance se déplaça néanmoins du charbon vers le pétrole et le gaz naturel. En 1900, l'énergie fournie par le charbon aux États-Unis valait dix fois celle fournie par le pétrole et le gaz naturel ensemble. Un demi-siècle plus tard, l'énergie fournie par le charbon ne représentait plus qu'un tiers de celle venant du pétrole et du gaz naturel.

Dans les temps anciens, l'huile utilisée dans les lampes pour l'éclairage venait de sources végétales et animales. Durant les éternités des temps géologiques, cependant, les minuscules animaux riches en huile des mers profondes ont parfois, en mourant, échappé à leurs prédateurs pour être enterrés, mélangés à la boue, sous des couches sédimentaires. Après une lente évolution chimique, l'huile a été transformée en un mélange complexe d'hydrocarbures appelé aujourd'hui *pétrole* (du mot latin signifiant « huile de roche »).

Le pétrole est parfois trouvé à la surface de la Terre, particulièrement dans le Moyen-Orient, qui en est si riche. C'était le *bitume* que Noé avait reçu l'ordre d'appliquer sur l'intérieur et l'extérieur de son arche pour la rendre étanche. De même, quand Moïse avait été abandonné enfant dans un berceau d'osier, ce dernier était enduit de bitume pour l'empêcher de

couler. De plus légères fractions du pétrole (*naphte*) étaient parfois recueillies et utilisées dans les lampes ou pour alimenter les flammes lors des rites religieux.

Vers 1850, on avait besoin de liquides inflammables pour les lampes. Il existait l'huile de baleine et l'huile de charbon (obtenue en chauffant du charbon en l'absence d'air). Une autre source était le schiste bitumineux, minerai mou qui ressemble un peu à de la cire. Quand il est chauffé, il fournit un liquide appelé *kérosène*. On en avait trouvé dans l'Ouest de la Pennsylvanie ; et en 1859, un conducteur de trains américain, Edwin Laurentine Drake, essaya quelque chose de nouveau.

Drake savait que l'on creusait des puits pour obtenir de l'eau, et que parfois on creusait très profond pour obtenir de la *saumure* (eau très salée dont on peut extraire du sel). Il venait quelquefois avec la saumure un produit huileux inflammable. Des récits rapportaient qu'en Chine et en Birmanie, deux mille ans plus tôt, on brûlait cette huile, et que la chaleur dégagée servait à faire évaporer l'eau de la saumure pour en extraire le sel.

Pourquoi donc ne pas creuser pour trouver cette huile ? On l'avait aussi utilisée, à cette époque ancienne, non seulement comme combustible pour les lampes, mais également à des fins médicales ; Drake sentait qu'il y aurait un marché pour ce qu'il pourrait extraire du sol. Il fora un trou de 21 mètres de profondeur à Titusville dans l'Ouest de la Pennsylvanie et, le 28 août 1859, il « tomba sur le pétrole ». Il avait foré le premier *puits de pétrole*.

Durant les cinquante années qui suivirent, les utilisations du pétrole furent limitées ; mais avec l'avènement du moteur à combustion interne, ce carburant devint très recherché. Une fraction liquide, plus légère que le kérosène (c'est-à-dire plus volatile et plus facilement convertie en vapeur), était précisément le carburant nécessaire aux nouveaux moteurs. La fraction était de l'*essence* ; la grande course au pétrole avait commencé.

Les champs de pétrole de Pennsylvanie furent vite épuisés, mais des champs beaucoup plus importants furent découverts au Texas au début du XXe siècle ; puis d'autres, plus importants encore, au Moyen-Orient au milieu du XXe siècle.

Le pétrole a beaucoup d'avantages sur le charbon. On n'a pas à descendre dans le sous-sol pour l'extraire ; on n'a pas à en charger d'innombrables camions ; on n'a pas à le stocker dans des caves ni à le jeter à la pelle dans les fourneaux ; il n'y a pas non plus de cendres à nettoyer. Le pétrole est pompé hors du sol, distribué par des tuyaux (ou par des pétroliers à travers les mers), stocké dans des réservoirs souterrains, et envoyé automatiquement dans les foyers, avec des flammes que l'on peut allumer et éteindre à volonté, et sans laisser aucune cendre. Surtout après la Deuxième Guerre mondiale, la planète tout entière eut tendance à abandonner le charbon pour le pétrole. Le charbon resta un matériau indispensable pour la fabrication du fer et de l'acier et pour d'autres applications, mais le pétrole devint la grande ressource du globe en carburant.

Le pétrole contient des fractions si volatiles qu'elles sont à l'état de vapeurs aux températures ordinaires. Elles constituent le *gaz naturel*. Le gaz naturel est encore plus facile à utiliser que le pétrole, et son usage s'est répandu encore plus rapidement que celui des fractions liquides du pétrole.

Pourtant, ces ressources sont limitées. Le gaz naturel, le pétrole et le charbon sont des *combustibles fossiles*, reliques de la vie animale et végétale qui florissait, il y a des éternités ; ils ne peuvent être remplacés au fur et à mesure de leur utilisation. Les hommes sont en train de consommer leur capital en combustibles fossiles à une vitesse extravagante.

Le pétrole, en particulier, s'en va fort vite en fumée. On brûle actuellement sur Terre plus de 636 millions de litres de pétrole à l'heure ; en dépit de tous les efforts d'économies d'énergie, ce taux de consommation continuera à augmenter dans le proche avenir. Bien qu'il en subsiste près de mille milliards de barils (1 baril = 159 litres) au sein de la Terre, cela représente au plus trente ans de consommation mondiale au rythme actuel.

Bien sûr, on peut fabriquer du pétrole synthétique en combinant du charbon ordinaire avec de l'hydrogène sous pression. Ce procédé fut mis au point pour la première fois par le chimiste allemand Friedrich Bergius dans les années vingt, ce qui lui valut de partager le prix Nobel de chimie en 1931. Les réserves de charbon sont énormes, de l'ordre de sept mille milliards de tonnes ; mais le charbon n'est pas toujours facile à extraire. Vers le XXVe siècle ou même avant, le charbon risque de devenir une ressource onéreuse.

Nous pouvons espérer découvrir de nouveaux gisements. Il est possible que nous attendent d'importantes surprises, sous forme de pétrole et de charbon, en Australie, au Sahara et même dans l'Antarctique. En outre, des perfectionnements techniques rendront peut-être moins coûteuses l'exploitation de filons de charbon plus étroits et plus profonds, celle de gisements de pétrole de plus en plus profonds et l'extraction du pétrole des schistes bitumineux et des réserves sous-marines.

Il ne fait aucun doute non plus que nous trouverons des moyens d'utiliser nos combustibles avec un meilleur rendement. Le procédé consistant à brûler du combustible pour produire de la chaleur, à utiliser cette chaleur pour convertir de l'eau en vapeur, et à faire entraîner par cette vapeur un générateur qui fournit de l'électricité, perd beaucoup d'énergie tout au long de la chaîne. Nombre de ces pertes pourraient être supprimées si l'on pouvait convertir directement la chaleur en électricité. Dès 1823, cela parut possible quand un physicien allemand, Thomas Johann Seebeck, observa que, si on reliait deux métaux différents en circuit fermé et que l'on chauffait la jonction des deux éléments, une aiguille de compas voisine du circuit était déviée, signe que la chaleur faisait passer un courant électrique dans le circuit (*thermoélectricité*). Seebeck n'avait pas interprété correctement sa propre expérience, cependant, et sa découverte n'eut pas de suites.

Avec l'arrivée des techniques de semi-conducteurs, le vieil *effet Seebeck* connut une renaissance. Les dispositifs thermoélectriques courants font usage de semi-conducteurs. Chauffer une extrémité d'un semi-conducteur crée un potentiel électrique dans le matériau ; dans un semi-conducteur de type p, l'extrémité froide devient négative, dans un type n elle devient positive. Si on relie ces deux types de semi-conducteurs en une structure en U avec la jonction n-p au bas du U, chauffer le bas rendra l'extrémité supérieure de la branche p négativement chargée et l'extrémité supérieure de la branche n chargée positivement. Il en résultera un courant qui circulera d'une extrémité à l'autre et qui subsistera tant que la différence de température sera maintenue (figure 10.1). (A l'inverse, l'injonction d'un courant peut entraîner une chute de température, si bien qu'un dispositif thermoélectrique peut être utilisé comme réfrigérateur.)

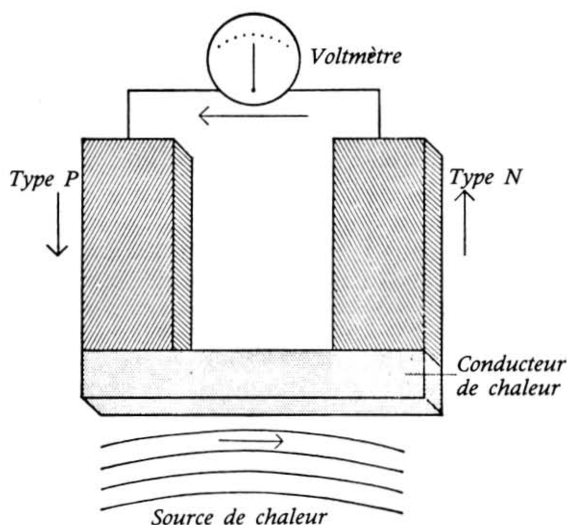


Figure 10.1. Elément thermoélectrique. Le chauffage du conducteur pousse les électrons à circuler vers l'extrémité froide du semi-conducteur de type n, et de la région froide à la région chaude du semi-conducteur de type p. Si on ferme le circuit, du courant circule dans le sens des flèches. De la chaleur est donc convertie en énergie électrique.

L'élément thermoélectrique, ne demandant ni générateur coûteux ni machine à vapeur encombrante, est portatif, et il peut être utilisé dans des endroits isolés comme générateur d'électricité à petite échelle. Tout ce qu'il lui faut comme source d'énergie est un brûleur au kérosène. On utilise de façon tout à fait courante de tels dispositifs dans des régions rurales de l'Union Soviétique.

En dépit de toutes les augmentations possibles du rendement d'utilisation des combustibles et des probabilités de nouvelles découvertes de gisements, tant de pétrole que de charbon, ces sources d'énergie sont forcément limitées. Le jour viendra, et dans peu de temps, où ni le charbon ni le pétrole ne pourront plus servir de sources d'énergie à grande échelle.

L'utilisation des combustibles fossiles devra diminuer, selon toute probabilité bien avant que les réserves actuelles soient épuisées, car l'augmentation de leur utilisation présente des dangers. La combustion des combustibles fossiles (particulièrement du charbon) dégage des oxydes d'azote et de soufre dans l'air. Une tonne de charbon n'en dégage pas trop ; mais avec toutes les combustions qui sont effectuées, c'est quelque 90 millions de tonnes d'oxydes de soufre qui ont été introduites dans l'atmosphère chaque année au cours des années soixante-dix.

De telles impuretés constituent l'une des sources essentielles de pollution de l'air et, dans certaines conditions météorologiques, sont à l'origine du *smog,* ce brouillard fumeux qui recouvre les villes, endommage les poumons et peut même tuer les personnes qui ont déjà des problèmes pulmonaires.

La pluie débarrasse l'air de cette pollution, mais elle crée aussi un autre problème, pire encore. Les oxydes d'azote et de soufre, en se dissolvant dans l'eau, la rendent très légèrement acide, de sorte qu'il tombe sur le sol de la *pluie acide.* La pluie n'est pas assez acide pour nuire à l'homme directement, mais elle tombe dans des mares et des lacs qu'elle acidifie – légèrement, mais assez pour tuer de nombreux poissons et autres êtres aquatiques, surtout si les lacs n'ont pas un fond calcaire qui peut partiellement neutraliser l'acide. Les pluies acides endommagent aussi les arbres. Les dégâts sont pires aux endroits où il y a une plus grande combustion de charbon et où la pluie tombe vers l'est en raison des vents d'ouest prédominants. Le Canada souffre donc des pluies acides dues à la combustion du charbon dans le Middle West américain, tandis que les Suédois subissent les contrecoups de la combustion de charbon de l'Europe de l'Ouest.

Les dangers d'une telle pollution peuvent devenir très importants si l'on continue à brûler les combustibles fossiles, et dans des proportions toujours croissantes. Des conférences internationales sont organisées à ce sujet. Pour améliorer la situation, il faudrait nettoyer le pétrole et le charbon avant de les brûler – procédé possible, mais qui ajouterait évidemment aux dépenses. Cependant, même si l'on ne brûlait comme charbon que du carbone pur et comme pétrole que des hydrocarbures purs, le problème persisterait. Le carbone, en brûlant, dégage du dioxyde de carbone, tandis que les hydrocarbures donnent du dioxyde de carbone et de l'eau. Ils sont relativement sans danger par eux-mêmes (bien qu'il puisse se former aussi du monoxyde de carbone, qui est fort toxique), et pourtant on ne peut éluder le problème.

Tant le dioxyde de carbone que la vapeur d'eau constituent des composants naturels de l'atmosphère. La quantité de vapeur d'eau varie selon l'endroit et le moment, mais le dioxyde de carbone est présent dans une proportion constante de l'ordre de 0,03 % en poids dans l'air. L'eau qui vient s'ajouter dans l'atmosphère par la combustion des combustibles fossiles trouve finalement son chemin vers l'Océan et ne représente qu'une variation insignifiante. Le dioxyde de carbone supplémentaire se dissout en partie dans les océans et réagit en partie avec les roches, mais il en reste dans l'atmosphère.

La quantité de dioxyde de carbone dans l'atmosphère a augmenté d'un facteur 1,5 par rapport à sa valeur en 1900, en raison de la combustion du charbon et du pétrole, et elle augmente d'année en année de façon mesurable. L'excès de dioxyde de carbone ne crée pas de problème au niveau de la respiration et peut même être considéré comme bénéfique pour la vie végétale. Cependant, il augmente quelque peu l'effet de serre de l'atmosphère et accroît légèrement la température moyenne totale de la Terre. Ceci est en soi à peine observable, mais l'augmentation de température tend à augmenter la pression de la vapeur de l'Océan, par conséquent à augmenter la teneur de l'air en vapeur d'eau, ce qui renforce également l'effet de serre.

Il est donc possible que la combustion des carburants fossiles déclenche une élévation de la température assez forte pour commencer à faire fondre les calottes glaciaires, ce qui aurait des résultats désastreux pour la ligne de côte des continents. Il peut aussi en résulter, au pire, des changements climatiques de grande envergure. Il existe même une faible possibilité pour que cela déclenche un renforcement de l'effet de serre tel que l'atmosphère

de la Terre se mettrait à ressembler à celle de Vénus ; mais nous avons besoin d'en apprendre beaucoup sur la dynamique de l'atmosphère et les effets de la température avant que les prédictions que nous faisons dépassent le stade des hypothèses.

En tout cas, le fait de continuer à brûler les combustibles fossiles doit être considéré avec infiniment de sérieux.

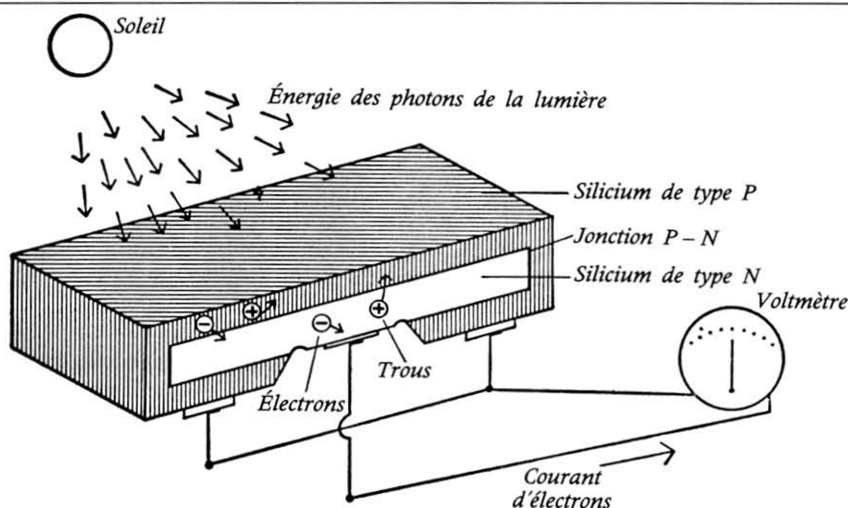
Pourtant, nos besoins en énergie existeront toujours et deviendront même supérieurs à ceux d'aujourd'hui. Que peut-on faire ?

L'ÉNERGIE SOLAIRE

Une possibilité consiste à faire un usage de plus en plus grand de sources d'énergie renouvelables, c'est-à-dire à vivre sur le revenu en énergie de la Terre plutôt que sur son capital. Le bois peut constituer une telle ressource si les forêts sont plantées et exploitées comme une culture, bien que le bois seul ne puisse arriver à satisfaire tous nos besoins en énergie. On peut en dire autant de certaines autres sources potentielles d'énergie sur la Terre, comme l'exploitation de la géothermie (dans les geysers, par exemple) ou l'utilisation des marées.

Bien plus importante, à long terme, est la possibilité de capter directement un peu de l'immense énergie déversée sur la Terre par le Soleil. L'*ensoleillement* produit de l'énergie à un taux environ 50 000 fois supérieur à notre taux moyen de consommation d'énergie. A cet égard, un dispositif particulièrement prometteur est la *cellule solaire* ou *cellule photovoltaïque,* qui utilise les dispositifs à l'état solide pour convertir directement en électricité la lumière du Soleil (figure 10.2).

Figure 10.2. Une cellule solaire. La lumière du Soleil arrivant sur la mince plaquette libère des électrons, formant ainsi des paires électrons-trous. La jonction p-n agit comme une barrière de champ électrique pour séparer les électrons des trous. Une différence de potentiel se développe donc à travers la jonction, et du courant circule à travers le circuit.



La cellule photovoltaïque mise au point en 1954 par les Laboratoires Bell est un sandwich plat de semi-conducteurs de type n et de type p, qui fait partie d'un circuit électrique. La lumière du Soleil heurtant la plaque arrache de leur place quelques électrons – l'effet photo-électrique ordinaire. Les électrons libérés se déplacent vers le pôle positif et les trous vers le pôle négatif, ce qui constitue un courant électrique. Le courant produit est faible par rapport à celui fourni par une pile chimique ordinaire, mais la grande réussite de la cellule solaire est qu'elle ne contient aucun liquide, aucun produit chimique corrosif, aucune partie en mouvement ; elle génère simplement de l'électricité, et cela indéfiniment tant qu'elle est exposée au soleil.

Le satellite artificiel *Vanguard I*, lancé par les États-Unis le 17 mars 1958, était le premier à être équipé d'une cellule photovoltaïque pour alimenter ses émetteurs radio ; ces signaux durèrent des années, puisqu'il n'existait pas d'interrupteur d'arrêt.

La quantité d'énergie tombant sur un hectare d'une région généralement ensoleillée est de 23 millions de kilowatt-heures par an. Si des surfaces importantes des régions désertiques de la Terre, comme la vallée de la Mort aux États-Unis ou le Sahara, étaient couvertes de cellules solaires et de dispositifs stockant l'électricité, elles pourraient satisfaire les besoins du monde en électricité pour un temps infini, aussi longtemps que persisterait l'espèce humaine.

Malheureusement, cela coûterait très cher. Les cristaux de silicium pur, dans lesquels il faut découper des tranches minces pour construire les cellules, sont fort coûteux. Pour fixer les idées, depuis 1958, si le prix des cellules n'est plus que 1/250 de son montant à l'origine, l'électricité solaire reste encore dix fois plus onéreuse que l'électricité produite à partir du pétrole.

Bien sûr, les cellules photovoltaïques peuvent devenir encore moins chères et de meilleur rendement, mais collecter la lumière du Soleil n'est pas aussi facile qu'il y paraît. La lumière solaire est abondante mais diffuse et, comme je l'ai mentionné plus haut, de vastes surfaces doivent être recouvertes de cellules solaires, si l'on veut qu'elles alimentent le monde. Ensuite, il fait nuit la moitié du temps, et pendant le jour il peut y avoir du brouillard, de la brume ou des nuages. Même l'air clair du désert absorbe une fraction non négligeable du rayonnement solaire, surtout quand le Soleil est bas dans le ciel. Enfin, l'entretien de ces grandes surfaces serait coûteux et difficile.

Certains scientifiques ont suggéré que des stations d'énergie solaire soient placées en orbite autour de la Terre dans des conditions telles que la lumière solaire soit presque ininterrompue et qu'il n'y ait pas d'absorption atmosphérique ; ce système pourrait augmenter la production par unité de surface par un facteur de 60, mais il est peu plausible que ce projet soit réalisé dans un avenir immédiat.

Le nucléaire et la guerre

Entre l'utilisation à grande échelle des combustibles fossiles à l'époque actuelle et celle de l'énergie solaire dans le futur, il y a une autre source

d'énergie disponible en grande quantité. Elle a fait son apparition de façon plutôt inattendue, il y a moins d'un demi-siècle, et elle pourrait faire la soudure. C'est l'*énergie nucléaire*, l'énergie emmagasinée dans le minuscule noyau de l'atome.

L'énergie nucléaire est parfois appelée *énergie atomique*, mais c'est une appellation erronée. A proprement parler, l'énergie atomique est l'énergie fournie par les réactions chimiques, comme par exemple la combustion du charbon et du pétrole, parce que ces réactions mettent en jeu le comportement de l'atome comme un tout. L'énergie libérée par des changements dans le noyau est d'une espèce totalement différente, et beaucoup plus forte.

LA DÉCOUVERTE DE LA FISSION

Peu après la découverte du neutron par Chadwick en 1932, les physiciens comprirent qu'ils possédaient un merveilleux outil pour explorer le noyau atomique. Comme il n'a pas de charge électrique, le neutron peut facilement pénétrer dans le noyau chargé. Les physiciens se mirent à bombarder divers noyaux avec des neutrons pour voir quelles réactions nucléaires ils pouvaient déclencher ; l'un des plus fervents utilisateurs du nouvel outil fut l'Italien Enrico Fermi. En l'espace de quelques mois, il avait préparé de nouveaux isotopes radioactifs de trente-sept éléments.

Fermi et ses collaborateurs découvrirent qu'ils obtenaient de meilleurs résultats s'ils ralentissaient les neutrons en les faisant d'abord passer à travers de l'eau ou de la paraffine. Par collision contre les protons de l'eau ou de la paraffine, les neutrons sont ralentis comme une balle de billard l'est quand elle en heurte d'autres. Quand la vitesse d'un neutron est réduite à la *vitesse thermique* (la vitesse normale du mouvement des atomes), il a une plus grande chance d'être absorbé par un noyau, parce qu'il reste plus longtemps au voisinage de celui-ci. Une autre explication du phénomène est que la longueur d'onde associée au neutron est alors plus grande, puisque la longueur d'onde est inversement proportionnelle à la quantité de mouvement de la particule. Quand le neutron ralentit, sa longueur d'onde augmente. Pour employer une image, le neutron devient plus « échevelé » et occupe plus de volume. Il heurte donc un noyau plus facilement, de même qu'une boule a plus de chances d'atteindre une quille qu'une balle de golf.

La probabilité qu'une espèce donnée de noyau capture un neutron s'appelle sa *section efficace*. Ce terme évoque, métaphoriquement, le noyau comme une cible de taille donnée. Il est plus facile d'atteindre avec une balle de tennis le mur d'une grange qu'un tableau noir situé à la même distance. Les sections efficaces des noyaux exposés au bombardement par des neutrons s'expriment en mille milliards de milliardièmes de centimètre carré (10^{-24} centimètre carré). Cette unité a été nommée *barn* par les physiciens américains M.G. Holloway et C.P. Baker en 1942.

Quand un noyau absorbe un neutron, son numéro atomique est inchangé (parce que la charge du noyau reste la même), mais son nombre de masse augmente d'une unité. L'hydrogène 1 devient l'hydrogène 2, l'oxygène 17 devient l'oxygène 18 et ainsi de suite. L'énergie fournie au noyau par le neutron quand il le pénètre peut *exciter* le noyau – c'est-à-dire augmenter l'énergie qu'il contient. Ce surplus d'énergie est alors émis sous forme de rayonnement gamma.

Le nouveau noyau est souvent instable. Par exemple, quand de l'aluminium 27 capture un neutron et devient de l'aluminium 28, l'un des neutrons dans le nouveau noyau se transforme bientôt en proton (en émettant un électron). Cette augmentation de la charge positive du noyau transforme l'aluminium (numéro atomique 13) en silicium (numéro atomique 14).

Comme le bombardement par neutrons fournit un moyen aisé pour convertir un élément en élément de numéro atomique immédiatement supérieur, Fermi décida de bombarder de l'uranium pour voir s'il pouvait se former un élément arficiel, le numéro 93. Dans les produits du bombardement de l'uranium, Fermi et ses collaborateurs trouvèrent des signes de nouvelles substances radioactives. Ils pensaient qu'ils avaient obtenu l'élément 93 et l'appelèrent *uranium X.* Mais comment le nouvel élément pouvait-il être identifié avec certitude ? Quelles sortes de propriétés chimiques devait-il avoir ?

L'élément 93, pensait-on, devait se trouver au-dessous du rhénium dans le tableau périodique ; il devait donc être chimiquement similaire au rhénium. (En vérité, personne ne savait à cette époque que l'élément 93 appartenait à une nouvelle série de terres rares, ce qui entraînait qu'il devait ressembler à l'uranium et non au rhénium – voir le chapitre 6. Les recherches en vue de son identification sont donc parties complètement dans la mauvaise direction.)

Si l'élément 93 ressemblait au rhénium, on aurait peut-être pu identifier les minuscules quantités qu'on avait créées en mélangeant les produits du bombardement par neutrons avec du rhénium, puis en séparant celui-ci par des méthodes chimiques. Le rhénium aurait agi comme un *porteur,* gardant avec lui l'« élément 93 » qui lui était chimiquement similaire. Si le rhénium avait alors présenté de la radioactivité, cela aurait indiqué la présence de l'élément 93.

Otto Hahn et Lise Meitner, qui avaient découvert le protoactinium en travaillant ensemble à Berlin, poursuivirent ce type d'expériences. Mais l'élément 93 ne fut pas détecté au moyen du rhénium. Hahn et Meitner continuèrent à essayer de trouver si le bombardement par neutrons n'avait pas transformé l'uranium en d'autres éléments proches de lui dans le tableau périodique. A ce moment, en 1938, l'Allemagne occupa l'Autriche et Meitner qui, jusque-là, en tant que citoyenne autrichienne, n'avait pas été inquiétée en dépit du fait qu'elle était juive, fut forcée de fuir l'Allemagne d'Hitler et partit à Stockholm. Hahn continua son travail avec le physicien allemand Fritz Strassmann.

Plusieurs mois plus tard, Hahn et Strassmann découvrirent que, quand on ajoutait à de l'uranium bombardé du baryum, celui-ci présentait de la radioactivité. Ils décidèrent que cette radioactivité devait être attribuée au radium, l'élément au-dessous du baryum dans le tableau périodique. Ils en conclurent donc que le bombardement de l'uranium par des neutrons en changeait une partie en radium. Mais ce radium semblait d'une espèce un peu spéciale. Quelle que soit la façon dont ils s'y prenaient, Hahn et Strassmann ne pouvaient le séparer du baryum. En France, Irène Joliot-Curie et son collaborateur P. Savitch entreprirent le même travail et échouèrent également. C'est alors que Meitner, réfugiée en Scandinavie, trancha l'énigme avec audace et publia une idée que Hahn avait émise en privé mais hésité à publier. Dans une lettre parue dans la revue anglaise *Nature* en janvier 1939, elle suggérait que le « radium » ne pouvait être

séparé du baryum parce que ce n'en était pas. Le radium supposé était en fait du baryum radioactif ; c'était du baryum qui s'était formé lors du bombardement de l'uranium par les neutrons. Ce baryum radioactif se désintégrait en émettant une particule bêta et formait du lanthanum. (Hahn et Strassmann avaient trouvé que du lanthanum ordinaire ajouté aux produits du bombardement par neutrons produisait de la radioactivité, qu'ils attribuèrent à de l'actinium ; en fait, il s'agissait de lanthanum radioactif.)

Mais comment former du baryum à partir de l'uranium ? Le baryum n'était qu'un atome de poids atomique moyen. Aucun processus connu de désintégration radioactive ne pouvait transformer un élément lourd en un autre dont le poids atomique ne vaudrait que la moitié. Meitner poussa l'audace jusqu'à suggérer que le noyau d'uranium s'était coupé en deux. L'absorption d'un neutron l'avait induit à subir ce qu'elle appelait une *fission*. Les deux éléments résultant de cette fission, disait-elle, étaient du baryum et l'élément 43, l'élément au-dessus du rhénium dans le tableau périodique. Un noyau de baryum et un noyau de l'élément 43 (appelé plus tard *technétium*) pouvaient reconstituer un noyau d'uranium. Ce qui en faisait une suggestion particulièrement audacieuse était le fait que le bombardement par neutrons n'apportait que six millions d'électron-volts, alors que, selon l'opinion scientifique de l'époque concernant la structure nucléaire, il en fallait plusieurs centaines de millions.

Un neveu de Meitner, Otto Robert Frisch, se précipita au Danemark pour présenter la nouvelle théorie à Bohr, avant même sa publication. Pour Bohr, il s'agissait de comprendre ce fait surprenant : le noyau se scindait facilement en deux. Heureusement, Bohr était justement en train de mettre au point une théorie de la structure nucléaire basée sur un modèle de goutte, et il lui sembla que sa théorie pouvait expliquer cette proposition surprenante. (Bien des années plus tard, le modèle de la goutte, en tenant compte des couches du noyau, permit d'expliquer les détails plus subtils de la fission et les raisons pour lesquelles le noyau se casse en parties inégales.)

De toute façon, quelle qu'en fût la théorie, Bohr saisit pleinement ce qu'impliquait la suggestion. Il devait justement aller à Washington assister à une conférence de physique théorique, et il y expliqua aux physiciens la proposition sur la fission qu'il avait entendue au Danemark. Très excités, les physiciens retournèrent à leurs laboratoires pour tester l'hypothèse, et en moins d'un mois, on annonça une demi-douzaine de confirmations expérimentales. Hahn obtint en 1944 le prix Nobel de chimie.

LA RÉACTION EN CHAÎNE

La réaction de fission libérait une quantité inhabituelle d'énergie, bien plus que la radioactivité ordinaire. Mais ce n'était pas seulement cette énergie supplémentaire qui rendait si extraordinaire le phénomène de la fission. Fait bien plus important, elle libérait deux ou trois neutrons. Dans les deux mois qui suivirent la publication de la lettre de Meitner, nombre de physiciens avaient envisagé la redoutable possibilité d'une *réaction nucléaire en chaîne.*

Une réaction en chaîne est un phénomène commun en chimie. La combustion d'un morceau de papier est un exemple de réaction en chaîne.

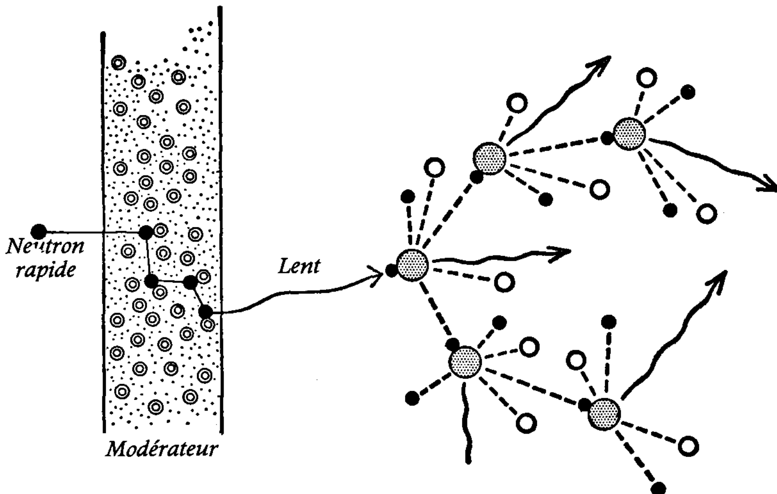
Une allumette fournit la chaleur nécessaire pour la faire démarrer ; une fois la combustion commencée, elle fournit l'agent même, la chaleur nécessaire pour maintenir et propager la flamme. La combustion entraîne plus de combustion qui, à son tour...

Cela est tout à fait similaire à une réaction nucléaire en chaîne. Un neutron entraîne la fission d'un noyau d'uranium, libérant deux neutrons qui peuvent produire deux fissions qui libéreront quatre neutrons qui pourront produire quatre fissions, et ainsi de suite (figure 10.3). Le premier atome à subir la fission libère 200 Mev d'énergie ; l'étape suivante fournit 400 Mev, la suivante 800 Mev, la suivante 1 600 Mev, et ainsi de suite. Comme les étapes successives ont lieu à peu près tous les 50 000 milliardièmes de seconde, on voit que, dans une minuscule fraction de seconde, sera libérée une quantité fantastique d'énergie. (Le nombre moyen réel de neutrons produits par la fission est 2,47, aussi les choses vont encore plus vite que ne l'indique ce calcul simplifié.) La fission de 10 grammes d'uranium produit autant d'énergie que la combustion de 30 tonnes de charbon ou de 2 500 litres de pétrole. Utilisé à des fins pacifiques, l'uranium pourrait, théoriquement, régler tous les problèmes dus à l'épuisement de nos sources de combustibles fossiles et à l'augmentation de notre consommation d'énergie.

Mais la découverte de la fission survint juste avant que le monde se trouve plongé dans une guerre généralisée. La fission de 10 grammes d'uranium fournirait, estimaient les physiciens, autant de pouvoir explosif que 600 tonnes de TNT. La pensée d'une guerre où l'on emploie de telles armes était terrifiante, mais l'idée que l'Allemagne nazie ait pu maîtriser un tel explosif avant les Alliés l'était encore plus.

Le physicien américain d'origine hongroise Leo Szilard, qui pensait aux réactions nucléaires en chaîne depuis des années, envisagea l'avenir avec

Figure 10.3. Réaction nucléaire en chaîne dans l'uranium. Les ronds gris sont des noyaux d'uranium, les ronds noirs des neutrons, les flèches ondulées des rayons gamma et les petits cercles des fragments de produits de fission.



lucidité. Lui-même et deux autres physiciens américains d'origine hongroise, Eugene Wigner et Edward Teller, persuadèrent durant l'été de 1939 le doux et pacifique Einstein d'écrire une lettre au président Franklin Delano Roosevelt pour lui signaler les potentialités de la fission de l'uranium, et suggérer que tous les efforts soient faits pour mettre au point une arme basée sur ce principe avant que les nazis n'y réussissent de leur côté.

La lettre fut écrite le 2 août 1939 et reçue par le Président le 11 octobre. Entre ces deux dates, la Deuxième Guerre mondiale avait éclaté en Europe. A l'université Columbia, des physiciens se mirent à travailler sous la direction de Fermi, qui avait quitté l'Italie pour l'Amérique l'année précédente, dans le but de produire une fission durable dans une grosse masse d'uranium.

Finalement, le gouvernement des États-Unis lui-même prit des mesures à la suite de la lettre d'Einstein. Le 6 décembre 1941, le président Roosevelt (prenant un gros risque politique en cas d'échec) autorisa l'organisation d'un immense projet connu sous le nom délibérément non compromettant de projet Manhattan, dans le but de mettre au point une bombe atomique. Le lendemain, les Japonais attaquaient Pearl Harbor et les États-Unis se retrouvaient en guerre.

LA PREMIÈRE PILE ATOMIQUE

Comme on pouvait s'y attendre, la pratique suivait à grand-peine la théorie. Ce n'était pas si simple d'obtenir une réaction en chaîne avec l'uranium. D'abord, il fallait disposer d'une quantité importante d'uranium, raffiné suffisamment pour que les neutrons ne soient pas gâchés dans des absorptions par les impuretés. L'uranium est un élément assez commun dans l'écorce terrestre, avec un taux moyen de deux grammes environ par tonne de roche, ce qui le rend 400 fois plus commun que l'or. Mais on le trouve de façon éparpillée, et il existe peu d'endroits sur le globe où il soit présent sous forme de minerai, riche ou même présentant des concentrations raisonnables. En outre, avant 1939, l'uranium n'avait presque aucune utilité, et on n'avait mis au point aucune méthode pour sa purification. On avait jusque-là produit aux États-Unis moins de trente grammes de métal d'uranium.

Les laboratoires de l'université d'Iowa se mirent, sous la direction de Spedding, à travailler au problème de la purification par résines échangeuses d'ions (voir le chapitre 6) et, en 1942, on commença à produire du métal d'uranium raisonnablement pur.

Cela, cependant, n'était qu'une première étape. Il fallait ensuite casser l'uranium pour séparer sa fraction la plus fissible. L'isotope d'uranium 238 (U-238) a un nombre pair de protons (92) et un nombre pair de neutrons (146). Les noyaux avec des nombres pairs de nucléons sont plus stables que ceux avec des nombres impairs. L'autre isotope présent dans l'uranium naturel, l'uranium 235 (U-235), a un nombre impair de neutrons (143). Bohr avait prédit pour cette raison qu'il serait plus fissible que l'uranium 238. En 1940, une équipe de recherche isola, sous la direction du physicien américain John Ray Dunning, une petite quantité d'uranium 235 et montra que l'hypothèse de Bohr était exacte. L'U-238 ne donne lieu au phénomène de fission que s'il est bombardé par des

neutrons rapides d'énergie supérieure à un certain seuil, tandis que l'U-235 subit la fission en absorbant des neutrons d'énergie quelconque, allant jusqu'aux simples neutrons thermiques.

Le problème suivant se posait : dans l'uranium naturel ce n'est qu'un atome sur 140 qui est de l'U-235, le reste étant de l'U-238. Donc, la plupart des neutrons libérés par la fission de l'U-235 seraient capturés par des atomes d'U-238 sans produire de fission. Même si on bombardait l'uranium avec des neutrons assez rapides pour faire subir la fission à l'U-238, les neutrons libérés par la fission de l'U-238 ne seraient pas assez énergétiques pour déclencher la réaction en chaîne dans les atomes restants de cet isotope plus courant. En d'autres termes, la présence de l'U-238 ferait s'amortir et mourir la réaction en chaîne. Ce serait comme d'essayer de faire brûler des feuilles humides.

Il n'y avait donc rien d'autre à faire que d'essayer de séparer à grande échelle l'U-235 de l'U-238, ou au moins d'enlever assez d'U-238 pour obtenir un enrichissement important du contenu en U-235 du mélange. Les physiciens s'attaquèrent au problème par plusieurs méthodes, chacune d'elles n'offrant que de minces chances de succès. Celle qui donna finalement les meilleurs résultats fut la *diffusion gazeuse*. Elle continua à être employée, bien qu'effroyablement onéreuse, jusqu'en 1960. Un scientifique de République fédérale d'Allemagne mit alors au point une technique de séparation de U-235, bien plus économique, par *centrifugation* : les molécules les plus lourdes sont envoyées vers l'extérieur tandis que les plus légères, contenant l'U-235, restent à l'intérieur. Ce procédé rend la fabrication de bombes nucléaires économiquement abordable aux pays de faibles ressources, une conséquence qui n'était pas véritablement désirée.

L'atome d'uranium 235 est 1,3 % moins massif que l'atome d'uranium 238. Il en résulte que, si les atomes étaient sous la forme d'un gaz, ceux d'U-235 se déplaceraient légèrement plus vite que les atomes d'U-238 et que l'on pourrait donc les séparer, grâce à leur diffusion plus rapide, à travers une série de barrières filtrantes. Mais il fallait d'abord convertir l'uranium en gaz. L'une des seules façons de l'obtenir sous cette forme était de le faire réagir avec du fluor, pour donner de l'*hexafluorure d'uranium*, un liquide volatil composé d'un atome d'uranium et de six atomes de fluor. Dans ce composé, une molécule contenant de l'U-235 est plus légère de moins de 1 % qu'une molécule contenant de l'U-238, différence qui s'avère suffisante pour que la méthode fonctionne.

On forçait l'hexafluorure d'uranium à traverser sous pression des parois poreuses. A chaque paroi, les molécules contenant U-235 traversaient légèrement plus vite ; ainsi, à chaque passage à travers les parois successives, la concentration en U-235 augmentait. Pour obtenir des quantités raisonnables d'hexafluorure d'uranium 235 presque pur, il fallait des milliers de parois, mais on pouvait quand même obtenir des concentrations d'U-235 assez riches avec un nombre de parois beaucoup plus petit.

Vers 1942, il apparaissait comme certain que la méthode de diffusion gazeuse (et une ou deux autres) pouvait produire de l'*uranium enrichi* en quantité ; on construisait en grand secret des usines de séparation (coûtant chacune un milliard de dollars et consommant autant d'électricité que toute la ville de New York) à Oak Ridge, dans le Tennessee.

Pendant ce temps, les physiciens calculaient la *masse critique* qui serait nécessaire pour maintenir une réaction en chaîne dans un fragment

d'uranium enrichi. Si le fragment était trop petit, un trop grand nombre de neutrons s'échapperaient de sa surface avant d'être absorbés par les atomes d'U-235. Pour diminuer cette perte, le volume du fragment devait être grand par rapport à sa surface. A une certaine masse critique, il y aurait assez de neutrons interceptés par des atomes d'U-235 pour maintenir la réaction en chaîne.

Les physiciens ont aussi trouvé une façon d'utiliser efficacement les neutrons disponibles. Les neutrons *thermiques* (c'est-à-dire lents) sont, comme je l'ai mentionné, plus facilement absorbés par l'U-235 que les neutrons rapides. Les expérimentateurs utilisaient donc un *modérateur* pour ralentir les neutrons, émis par la réaction de fission à des vitesses plutôt élevées. L'eau ordinaire aurait constitué un excellent modérateur, mais malheureusement, les noyaux d'hydrogène ordinaire absorbent avidement les neutrons. Le deutérium (hydrogène 2) remplit beaucoup mieux cette fonction ; il n'a pratiquement aucune tendance à absorber les neutrons. C'est pourquoi les expérimentateurs travaillant sur la fission s'intéressèrent de près à la préparation de réserves d'eau lourde.

Jusqu'en 1943, cette dernière fut surtout préparée par électrolyse. L'eau ordinaire se sépare en hydrogène et oxygène plus facilement que l'eau lourde, de sorte que, si on électrolysait une grande quantité d'eau, la quantité d'eau restante était riche en eau lourde et pouvait être conservée. Après 1943, on préféra une méthode basée sur une distillation poussée. L'eau ordinaire avait un point d'ébullition plus bas, de sorte que la dernière goutte d'eau non distillée était riche en eau lourde.

L'eau lourde avait paru extrêmement précieuse dès 1940. Une histoire rocambolesque raconte la façon dont Joliot-Curie parvint à soustraire à la barbe et au nez des envahisseurs nazis tout le stock français de ce liquide. Quelque quatre cents litres, qui avaient été préparés en Norvège, tombèrent aux mains des nazis allemands. Un commando britannique les détruisit dans un raid en 1942.

Pourtant, l'eau lourde présentait des inconvénients ; elle pouvait se mettre à bouillir quand la réaction en chaîne produirait de la chaleur et cela pouvait attaquer l'uranium.

Les scientifiques du projet Manhattan qui cherchaient à créer un système permettant la réaction en chaîne décidèrent d'utiliser du carbone, sous forme de graphite très pur, comme modérateur.

Un autre modérateur possible, le béryllium, avait l'inconvénient d'être toxique. D'ailleurs, l'affection connue sous le nom de *bérylliose* fut observée pour la première fois au début des années quarante chez l'un des physiciens travaillant sur la bombe atomique.

Imaginons maintenant une réaction en chaîne. Nous faisons démarrer la chaîne en envoyant un faisceau de neutrons de déclenchement dans l'ensemble constitué par le modérateur et l'uranium enrichi. Nombre d'atomes d'uranium 235 vont subir la fission en libérant des neutrons qui vont heurter d'autres atomes d'uranium 235. Ceux-ci subissent à leur tour la fission et libèrent d'autres neutrons. Quelques neutrons seront absorbés par d'autres atomes que ceux d'uranium 235 ; certains s'échapperont complètement de la pile. Mais si dans chaque fission il y a au moins un neutron qui produit une autre fission, alors la réaction en chaîne sera auto-entretenue. Si le *coefficient multiplicateur* est plus grand que 1, même très légèrement (par exemple 1,001), la réaction en chaîne conduira rapidement à une explosion. Ceci est bien adapté à la fabrication des

bombes, mais moins à la réalisation d'expériences. Pour faire des expériences, il faut un dispositif qui permette de contrôler le taux de fission. On y arrive en faisant pénétrer dans le « bain » où se fait la réaction des barres d'une substance comme le cadmium, qui a une grande section efficace pour la capture des neutrons. La réaction en chaîne se développe si rapidement que l'on ne pourrait faire glisser les barres de cadmium assez rapidement, si n'intervenait pas le fait que les atomes d'uranium en train de subir la fission n'émettent pas tous leurs neutrons instantanément. Près d'un neutron sur 150 est un *neutron retardé* émis quelques minutes après la fission, puisqu'il sort, non pas directement des atomes en train de subir la fission, mais de plus petits atomes formés durant la fission. Quand le coefficient de multiplication est à peine au-dessus de 1, ce délai est suffisant pour donner le temps d'appliquer les barres de contrôle.

En 1941, on fit des expériences sur des mélanges uranium-graphite et on rassembla assez d'informations pour que les physiciens décident que, même sans uranium enrichi, on pouvait mettre en œuvre une réaction en chaîne si seulement le fragment d'uranium était assez grand.

Les physiciens se mirent à construire un réacteur en chaîne à uranium de masse critique à l'université de Chicago. A cette époque on disposait à peu près de six tonnes d'uranium pur ; cette quantité était complétée par de l'oxyde d'uranium. On empila les unes sur les autres des couches alternées d'uranium et de graphite, cinquante-sept en tout, munies de trous pour insérer les barres de contrôle en cadmium. Cette structure s'appelait une *pile*, nom de code vague ne laissant pas transparaître l'usage qui en était prévu. (Durant la Première Guerre mondiale, c'est pour les mêmes raisons de secret que l'on appela *tanks* les nouveaux véhicules blindés sur chenilles. Le nom de tank s'est perpétué, mais celui de *pile atomique* a fini par être remplacé par le terme plus parlant de *réacteur nucléaire*.)

La pile de Chicago, construite sous le stade de football, mesurait 9 mètres de large, 9,75 mètres de long et 6,5 mètres de haut. Elle pesait 1 420 tonnes et contenait 53 tonnes d'uranium sous forme de métal et d'oxyde. (Si on avait pu utiliser de l'uranium 235 pur, la taille critique n'aurait pas dépassé 260 grammes.) Le 2 décembre 1942, on retira lentement les barres de contrôle en cadmium. A 15 h 45, le coefficient multiplicateur atteignait 1 : une fission auto-entretenue commençait. A ce moment-là l'humanité entra (sans le savoir) dans l'ère nucléaire.

Le physicien responsable était Enrico Fermi, et Eugene Wigner lui offrit une bouteille de Chianti. Arthur Compton, qui était présent, téléphona à James Bryant Conant à Harvard pour lui annoncer la nouvelle. « Le navigateur italien, dit-il, est arrivé au Nouveau Monde. » Conant demanda : « Comment étaient les indigènes ? » La réponse fut : « Très amicaux ! »

Il est curieux et intéressant de noter qu'un navigateur italien avait découvert un nouveau monde en 1492 et qu'un second en découvrait un autre en 1942.

L'ÈRE NUCLÉAIRE

Pendant ce temps un autre combustible utilisable pour la fission s'était révélé. L'uranium 238 forme, par absorption d'un neutron thermique, de l'uranium 239, qui se désintègre rapidement en neptunium 239, qui à son tour se désintègre presque aussi vite en plutonium 239.

Comme le noyau de plutonium 239 a un nombre impair de neutrons (145) et qu'il est plus complexe que l'uranium 235, il devait être fortement instable. On avait toutes les raisons de penser que le plutonium 239 donnerait lieu à la fission sous l'action de neutrons thermiques, comme l'uranium 235 ; ce fut confirmé expérimentalement en 1941. Comme ils n'étaient pas encore sûrs que la préparation de l'uranium 235 serait réalisable pratiquement, les physiciens avaient décidé de répartir les risques en essayant aussi de produire du plutonium en quantité.

Pour produire du plutonium, on construisit des réacteurs spéciaux en 1943 à Oak Ridge et à Hanford, dans l'État de Washington. Ces réacteurs représentaient un grand progrès par rapport à la première pile de Chicago. D'abord, les nouveaux réacteurs étaient conçus de façon à pouvoir en retirer périodiquement l'uranium. Le plutonium produit pouvait être séparé de l'uranium par des méthodes chimiques ; et les produits de fission, dont certains étaient d'énergiques absorbants de neutrons, pouvaient aussi en être séparés. En outre, les nouveaux réacteurs étaient refroidis à l'eau pour empêcher un échauffement excessif. (La pile de Chicago ne pouvait fonctionner que durant de courtes périodes, parce qu'elle était simplement refroidie à l'air.)

Vers 1945, on avait assez d'uranium 235 purifié et de plutonium 239 pour pouvoir construire des bombes. Cette partie de la tâche fut entreprise dans une troisième cité secrète, Los Alamos, au Nouveau-Mexique, sous la direction du physicien américain J. Robert Oppenheimer.

Dans le cas d'une bombe, il était nécessaire que la réaction en chaîne s'établisse le plus vite possible. Ceci conduisait à utiliser des neutrons rapides pour effectuer la réaction, et à raccourcir les intervalles entre les fissions, donc à supprimer le modérateur. La bombe était aussi enfermée dans une enceinte massive pour maintenir l'uranium en un seul bloc assez longtemps pour qu'une forte proportion en subisse la fission.

Comme une masse critique de matériau fissible aurait explosé spontanément (les quelques neutrons existant dans l'atmosphère suffisant à déclencher la réaction), le combustible de la bombe était divisé en deux blocs ou plus. Le mécanisme de mise à feu était constitué par un explosif qui rapprochait ces blocs quand il fallait faire exploser la bombe. Une des configurations adoptées était appelée « le maigre » – un tube muni à ses deux extrémités de deux blocs d'uranium 235. Une autre, « le gros », avait la forme d'une sphère dans laquelle une coquille composée de matériau fissible pouvait subir une *implosion* vers le centre, créant ainsi une masse critique dense maintenue momentanément par la force de l'implosion et par une lourde gaine extérieure appelée le *réflecteur*. Le réflecteur servait aussi à renvoyer des neutrons dans la masse en fission, abaissant ainsi la masse critique.

Essayer un tel dispositif à échelle réduite était impossible. La bombe devait être au-dessus de la taille critique ou ne pas être. Le premier essai fut donc l'explosion d'une bombe à fission nucléaire grandeur nature, appelée habituellement (et incorrectement) *bombe atomique* ou *bombe A*. A 5 h 30 du matin le 16 juillet 1945, à Alamogordo, au Nouveau-Mexique, une bombe explosa avec un effet vraiment terrifiant ; elle avait la puissance explosive de 20 000 tonnes de TNT. On dit que I.I. Rabi, à qui on demandait plus tard ce qu'il avait observé, répondit d'un air catastrophé : « Je ne peux vous le dire, mais ne vous attendez pas à mourir de mort naturelle »

(ajoutons que celui à qui s'adressait Rabi est mort de mort naturelle quelques années plus tard).

On prépara deux autres bombes à fission. L'une, une bombe à uranium surnommée « le petit garçon », de 3 mètres de long sur 60 centimètres de large et d'un poids de 4,5 tonnes, fut lancée sur Hiroshima le 6 août 1945 ; elle fut mise à feu par écho radar. Quelques jours plus tard, la seconde, une bombe au plutonium, de 3 mètres sur 1,5 mètre, pesant 5 tonnes et surnommée « le gros », fut larguée sur Nagasaki. Ensemble, les deux bombes représentaient une puissance explosive équivalente à celle de 35 000 tonnes de TNT. Avec le bombardement d'Hiroshima, l'ère nucléaire, déjà presque vieille de trois ans, secoua la conscience du monde.

Pendant les quatre années qui suivirent, les Américains vécurent dans l'illusion qu'il existait un « secret » de la bombe nucléaire que l'on pourrait taire aux autres nations à condition de prendre des mesures de sécurité suffisantes. En réalité, les faits et les théories de la fission nucléaire faisaient l'objet de publications depuis 1939, et l'Union Soviétique était bien engagée dans la recherche sur ce sujet dès 1940. Si la Seconde Guerre mondiale n'avait mobilisé les ressources de cette nation bien davantage que celles, beaucoup plus grandes, des États-Unis, qui n'avaient pas été envahis, l'Union Soviétique aurait pu fabriquer une bombe nucléaire en 1945. Elle fit exploser sa première bombe nucléaire le 22 septembre 1949, provoquant l'effroi et un étonnement bien injustifié chez la plupart des Américains. Sa puissance valait six fois celle de la bombe d'Hiroshima, et les effets de l'explosion égalaient ceux de 210 000 tonnes de TNT.

Le 3 octobre 1952, la Grande-Bretagne devenait la troisième puissance nucléaire en faisant exploser une bombe d'essai de sa fabrication. Le 13 février 1960, la France rejoignait le « club nucléaire » en qualité de quatrième membre, en mettant à feu une bombe au plutonium dans le Sahara. Le 16 octobre 1964, la république populaire de Chine annonça l'explosion d'une bombe nucléaire et devint le cinquième membre du club. En mai 1974, l'Inde fit exploser une bombe nucléaire, en utilisant du plutonium qui avait été secrètement extrait d'un réacteur (destiné à la production pacifique d'énergie) dont le Canada lui avait fait don, et elle devint ainsi le sixième membre. Depuis lors, nombre de puissances, parmi lesquelles Israël, l'Afrique du Sud, l'Argentine et l'Irak sont en passe, semble-t-il, de posséder elles aussi leurs armes nucléaires.

Une telle *prolifération d'armes nucléaires* est devenue un facteur de crainte pour beaucoup de gens. Il n'est déjà pas rassurant de vivre sous la menace d'une guerre nucléaire déclenchée par l'une des deux superpuissances. Celles-ci sont, malgré elles, conscientes des conséquences et elles se sont retenues pendant quarante ans de le faire. Etre en outre à la merci de petites puissances, agissant sous le coup de la colère et avec une très faible marge de manœuvre, paraît intolérable.

LA RÉACTION THERMONUCLÉAIRE

Pendant ce temps, la bombe à fission avait été largement dépassée. Les hommes avaient réussi à découvrir une autre réaction nucléaire énergétique qui rendait réalisables des bombes encore plus dévastatrices.

Dans la fission de l'uranium, 0,1 % de la masse de l'atome d'uranium est convertie en énergie. Mais dans la fusion des atomes d'hydrogène pour

former de l'hélium, c'est 0,5 % de leur masse qui est convertie en énergie ; en 1915, le chimiste américain William Draper Harkins l'avait remarqué le premier. A des températures de quelques millions de degrés, l'énergie des protons est assez élevée pour leur permettre de fondre. Deux protons peuvent donc s'unir et, après avoir émis un positron et un neutrino (un processus qui convertit l'un des protons en un neutron), devenir un noyau de deutérium. Un noyau de deutérium peut alors s'unir à un proton pour former un noyau de tritium, qui peut ensuite, toujours par fusion avec un autre proton, former un atome d'hélium 4. Ou encore les noyaux de deutérium et de tritium se combinent de façons diverses pour former de l'hélium 4.

Comme de telles réactions nucléaires ne se produisent que sous l'effet de hautes températures, on les qualifie de *réactions thermonucléaires.* Dans les années trente, on croyait que le seul endroit où pouvaient exister les températures nécessaires était le centre des étoiles. En 1938, le physicien allemand Hans Albrecht Bethe (qui avait fui l'Allemagne hitlérienne pour les États-Unis en 1935) émit l'idée que les réactions de fusion étaient responsables de l'énergie que rayonnaient les étoiles. C'était la première explication complètement satisfaisante de l'énergie stellaire depuis que Helmholtz avait soulevé la question près d'un siècle plus tôt.

Désormais, la bombe à fission à uranium fournissait les températures nécessaires sur la Terre. Elle pouvait servir d'allumette assez chaude pour déclencher une réaction en chaîne de fusion dans l'hydrogène. Pendant un certain temps, on ne crut pas vraiment que la réaction permettrait de faire une bombe. La raison en était que le combustible hydrogène, sous la forme d'un mélange de deutérium et de tritium, devait être condensé en une masse dense, ce qui voulait dire qu'il devait être liquéfié et maintenu à une température d'à peine quelques degrés au-dessus du zéro absolu. En d'autres termes, ce qui devrait exploser serait en fait un énorme réfrigérateur. En outre, à supposer qu'on puisse réaliser une bombe à hydrogène, quel usage en faire ? La bombe à fission était déjà assez dévastatrice pour détruire des villes ; une bombe à hydrogène aurait simplement multiplié les dégâts et anéanti des populations civiles entières. Néanmoins, malgré ces perspectives peu encourageantes, les États-Unis et l'Union Soviétique se sentirent obligés de continuer dans cette voie. La commission de l'énergie atomique des États-Unis se mit en mesure de produire du tritium, construisit un curieux engin à fission et à fusion sur un atoll corallien du Pacifique et, le 1er novembre 1952, fit exploser la première bombe thermonucléaire (une *bombe à hydrogène* ou *bombe H*) sur notre planète. Elle correspondait à toutes les prévisions : l'explosion libéra l'équivalent de 10 millions de tonnes de TNT (*10 mégatonnes*) – 500 fois la modeste énergie de 20 kilotonnes de la bombe d'Hiroshima. Le souffle détruisit l'atoll.

Les Russes n'étaient pas loin derrière ; le 12 août 1953, ils réussirent à obtenir une explosion thermonucléaire, et leur dispositif était assez léger pour être lancé d'un avion. Les États-Unis ne produisirent de version mobile qu'au début de 1954. Alors qu'ils avaient mis au point la bombe à fusion sept ans et demi après la bombe à fission, les Soviétiques n'y mirent que cinq ans.

Pendant ce temps on avait conçu une méthode détournée pour générer une réaction thermonucléaire en chaîne d'une manière plus simple et la monter dans une bombe transportable. La clé en était l'élément lithium.

Quand l'isotope 6 du lithium absorbe un neutron, il se scinde en noyaux d'hélium et de tritium en libérant 4,8 Mev d'énergie. Supposons, alors, qu'on utilise comme combustible un composé de lithium et d'hydrogène (ce dernier sous la forme de son isotope lourd, le deutérium). Ce composé étant un solide, il n'y a aucun besoin de réfrigération pour condenser le combustible. Un allumeur à fission fournirait des neutrons pour scinder le lithium, et la chaleur de l'explosion entraînerait la fusion du deutérium présent dans le composé et du tritium produit par la désintégration du lithium. En d'autres termes, il se produirait plusieurs réactions libérant de l'énergie : la fission du lithium, la fusion du deutérium avec le deutérium et la fusion du deutérium avec le tritium.

Mais à part la libération de ces gigantesques énergies, ces réactions fourniraient aussi un grand nombre de neutrons en excès. Les constructeurs de bombe se demandèrent alors s'il n'était pas possible d'utiliser ces neutrons pour déclencher la fission d'une masse d'uranium. Même l'uranium 238 ordinaire pourrait subir la fission du fait de neutrons rapides (bien moins facilement toutefois que l'U-235). L'intense émission de neutrons rapides due aux réactions de fusion peut entraîner la fission d'un nombre considérable d'atomes d'U-238. Supposons que l'on construise une bombe avec un noyau central d'U-235 (« l'allumette ») entouré d'une charge explosive d'hydrure lourd de lithium, avec pour finir une couche d'uranium 238 qui puisse aussi servir d'explosif. Cela ferait une bien grosse bombe. On pourrait donner à la couche d'U-238 presque autant d'épaisseur qu'on le voudrait parce qu'il n'y a pas de masse critique à partir de laquelle l'uranium 238 subit spontanément une réaction en chaîne. On appelle parfois cette construction la *bombe U.*

On a construit cette bombe. Elle a explosé à Bikini, dans les îles Marshall, le 1er mars 1954, à l'effroi du monde entier. L'énergie dégagée était d'environ 15 mégatonnes. Mais, bien plus dramatique, une pluie de particules radioactives tomba sur vingt-trois pêcheurs japonais du bateau *Le Dragon chanceux.* La radioactivité détruisit la cargaison de poisson, rendit malades les pêcheurs, en tuant même un, et n'améliora pas vraiment la santé du reste du monde.

Depuis 1954, les bombes thermonucléaires font partie des armements des États-Unis, de l'Union Soviétique et de la Grande-Bretagne. En 1967, la Chine est devenue le quatrième membre du « club thermonucléaire », après avoir mis trois ans seulement à passer de la fission à la fusion. L'Union Soviétique a fait exploser des bombes à hydrogène dans les 50 à 100 mégatonnes ; les États-Unis sont parfaitement capables de construire de telles bombes, ou même de plus grosses, à bref délai.

Dans les années soixante-dix, on a mis au point des bombes thermonucléaires en minimisant l'effet du souffle et en maximisant le rayonnement, particulièrement celui des neutrons, ce qui entraînerait moins de dégâts matériels, mais davantage de dommages pour les êtres humains. De telles *bombes à neutrons* semblent intéressantes à ceux qui accordent plus de prix aux biens matériels qu'aux vies humaines. Quand on avait largué les premières bombes nucléaires utilisées à la fin de la Seconde Guerre mondiale, c'est à l'aide d'avions qu'on l'avait fait. Il est maintenant possible de les envoyer sur l'objectif avec des *missiles balistiques intercontinentaux* (MBIC) qui sont propulsés par fusée et susceptibles d'être dirigés avec une très grande précision d'un point de la Terre à un autre. Les États-Unis et l'Union Soviétique possèdent des stocks importants de

missiles de ce genre, pouvant être éventuellement équipés de têtes nucléaires.

C'est pour cette raison qu'une guerre thermonucléaire entre les deux superpuissances, qui pourrait s'engager sous l'effet d'un accès de fureur insensé des deux côtés, mettrait un terme à la civilisation (et peut-être aussi à beaucoup des facultés qu'a la Terre d'entretenir la vie) en moins d'une demi-heure. C'est là, on en conviendra, une pensée bien affligeante.

L'utilisation pacifique du nucléaire

L'usage de l'énergie nucléaire sous la forme de bombes d'une incroyable puissance de destruction a contribué, plus que tout ce qui s'est passé depuis les débuts de la science, à faire du scientifique un véritable ogre. Dans un sens, ce portrait caricatural est justifié, car aucun argument, aucun raisonnement ne peut changer le fait que des scientifiques ont effectivement construit la bombe nucléaire, en connaissant dès le début sa puissance dévastatrice et en sachant qu'elle serait probablement utilisée.

Il faut ajouter qu'ils ont agi ainsi dans le contexte d'une guerre importante contre des ennemis sans pitié, avec à l'esprit l'éventualité terrifiante qu'un maniaque dangereux comme Adolphe Hitler puisse disposer le premier d'une telle bombe. On doit aussi ajouter que les scientifiques qui travaillaient sur la bombe étaient divisés à son propos ; beaucoup d'entre eux étaient opposés à son emploi, tandis que nombre d'autres ont quitté le domaine de la physique nucléaire après la guerre dans un mouvement de remords.

En 1945, un groupe de physiciens, conduit par le prix Nobel James Franck, a envoyé une pétition au secrétaire d'État à la Défense contre l'usage de la bombe nucléaire sur les villes japonaises, tout en prédisant avec justesse le dangereux équilibre de la terreur qui suivrait son emploi. On a remarqué beaucoup moins de crises de conscience chez les dirigeants politiques et militaires qui avaient pris la décision d'utiliser les bombes et qui, pour d'étranges raisons, sont considérés comme des patriotes par ceux-là mêmes qui voient dans les scientifiques des démons.

En outre, nous ne pouvons et ne devons pas sous-estimer le fait qu'en libérant l'énergie du noyau atomique, les scientifiques ont mis à notre disposition une énergie utilisable aussi bien à des fins constructives que destructives. Il est important d'insister sur ce point dans un monde et à une époque où la menace de la destruction nucléaire a placé la science et les scientifiques dans une position défensive un peu honteuse. Il importe que ceci soit dit et proclamé dans un pays comme les États-Unis où une forte tradition rousseauiste dénigre l'enseignement livresque, censé corrompre l'intégrité naturelle des êtres humains.

Même l'explosion d'une bombe atomique peut ne pas être purement destructrice. Comme les explosifs chimiques bien moins puissants utilisés depuis longtemps dans les mines et la construction des barrages et des autoroutes, les explosifs nucléaires pourraient être très utiles dans des projets de construction. On a rêvé à toutes sortes d'applications : creuser des ports, des canaux, faire sauter des formations rocheuses souterraines, préparer des réservoirs de chaleur pour en tirer l'énergie – et même

propulser des vaisseaux spatiaux à longue distance. Dans les années soixante, cependant, l'excitation suscitée par ces espoirs lointains s'est apaisée. En effet, les perspectives des dangers de contamination radioactive ou de dépenses excessives, ou les deux à la fois, ont agi comme un éteignoir.

Il y a pourtant une utilisation constructive de l'énergie nucléaire qui a été réalisée ; elle était en germe dans la réaction en chaîne qui avait vu le jour sous le stade de football de l'université de Chicago. Un réacteur nucléaire contrôlé peut fournir d'énormes quantités de chaleur que l'on peut extraire en utilisant un fluide caloporteur comme l'eau, ou même du métal fondu, pour produire de l'électricité ou de la chaleur (figure 10.4).

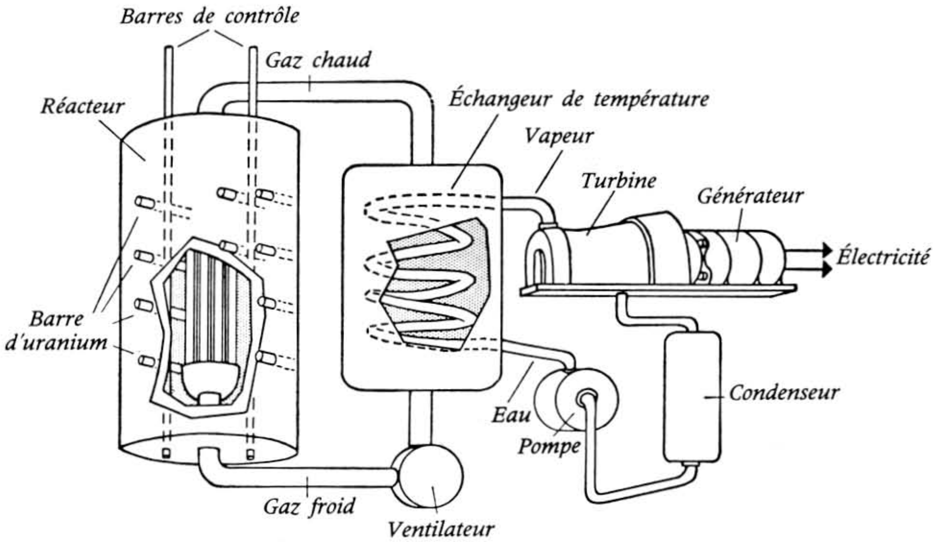


Figure 10.4. Centrale nucléaire du type à refroidissement à gaz, représentée par un schéma. La chaleur du réacteur est transférée à un gaz qui le traverse ; ce gaz peut être un métal vaporisé, et la chaleur sert à convertir de l'eau en vapeur.

LES NAVIRES À PROPULSION NUCLÉAIRE

Des réacteurs nucléaires qui produisaient de l'électricité ont été construits en Grande-Bretagne et aux États-Unis peu après la guerre. Les États-Unis possèdent à présent une flotte de plus de cent sous-marins à propulsion nucléaire, dont le premier, le *Nautilus* (d'un coût de 50 millions de dollars) a été lancé en janvier 1954. Ce navire, aussi important pour son époque que le *Clermont* de Fulton pour la sienne, avait des machines disposant d'une énergie presque illimitée, ce qui lui permettait de rester en plongée pendant des périodes d'une longueur indéfinie, alors que les submersibles ordinaires doivent souvent faire surface pour recharger leurs batteries à l'aide de générateurs diesel qui ont besoin d'air pour fonctionner. En outre,

alors que les sous-marins ordinaires atteignent des vitesses de huit nœuds, un sous-marin nucléaire fait du vingt nœuds ou plus.

Le premier cœur du réacteur du *Nautilus* a duré jusqu'à ce que le navire ait parcouru 54 300 milles marins. Le *Nautilus* a fait en 1958 une traversée sous-marine de l'océan Arctique. Ce voyage a permis d'établir que la profondeur de l'océan au pôle Nord était de 4 087 mètres (2,2 milles marins), c'est-à-dire supérieure aux estimations antérieures. Un second submersible nucléaire plus grand, le *Triton*, a fait le tour du globe en plongée en suivant la route de Magellan en 84 jours, de février à mai 1960.

L'Union Soviétique possède aussi des submersibles nucléaires ; elle a lancé en décembre 1957 le premier navire de surface à propulsion nucléaire, le *Lénine*, un brise-glace. Un peu avant, les États-Unis avaient mis en chantier un vaisseau de surface à propulsion nucléaire, et en juillet 1959, ils ont lancé le *Long Beach* (un croiseur) et le *Savannah* (un vaisseau marchand). Le *Long Beach* est propulsé par deux réacteurs nucléaires.

Moins de dix ans après le lancement des premiers navires à propulsion nucléaire, les États-Unis avaient quatre navires de surface à propulsion nucléaire, en opération, en construction ou en projet adopté. Et pourtant, l'enthousiasme pour la propulsion nucléaire faiblissait, sauf pour les sous-marins. En 1967, le *Savannah* fut retiré du service au bout de deux ans. Son entretien coûtait trois millions de dollars par an ; il était donc considéré comme trop onéreux.

LES CENTRALES ÉLECTRIQUES NUCLÉAIRES

Les militaires ne sont pas les seuls à devoir être servis. Le premier réacteur nucléaire construit pour la production d'électricité à usage civil fut mis en service en Union Soviétique en juin 1954. C'était une petite unité, d'une puissance de 5 000 kilowatts. En octobre 1956, la Grande-Bretagne démarrait sa centrale de Calder Hall, d'une capacité de plus de 50 000 kilowatts. Les États-Unis arrivèrent en troisième position. Le 26 mai 1958, Westinghouse mettait en service à Shippingport, en Pennsylvanie, un petit réacteur nucléaire pour la production d'énergie à usage civil d'une puissance de 60 000 kilowatts. D'autres réacteurs ont suivi assez vite, tant aux États-Unis qu'ailleurs.

En un peu plus de dix ans, il y avait des réacteurs nucléaires dans une douzaine de pays, et près de la moitié de l'électricité civile aux États-Unis était fournie par la fission des noyaux. Même l'espace interplanétaire était envahi, puisqu'un satellite alimenté par un petit réacteur fut lancé le 3 avril 1965. Pourtant, le problème de la contamination radioactive est très sérieux. Au début des années soixante-dix, l'opposition du public à la prolifération continue des centrales nucléaires s'est renforcée.

C'est alors que, le 28 mars 1979, à Three Mile Island, sur la rivière Susquehanna près d'Harrisburg, s'est produit le plus sérieux accident nucléaire de l'histoire américaine. En réalité, on n'a finalement décelé aucune augmentation de la radioactivité, aucun signe pouvant mettre en danger la vie humaine, bien que la panique ait régné pendant plusieurs jours. Le réacteur a été cependant mis définitivement hors service, et toute remise en état des lieux devrait être très longue et coûteuse. La principale victime a été l'industrie nucléaire de production d'énergie. Une vague de

sentiment antinucléaire a balayé les États-Unis ainsi que d'autres nations, et la probabilité de voir mis en route de nouveaux réacteurs nucléaires aux États-Unis s'est fortement estompée.

L'accident, s'il a placé les Américains face au danger de la contamination radioactive, a, semble-t-il, mobilisé également l'opinion mondiale contre la production (laissons de côté l'usage) des bombes nucléaires, résultat qui ne peut apparaître que comme positif à tout être doué de raison.

Pourtant, on ne peut pas si facilement renoncer à l'énergie nucléaire dans ses aspects pacifiques. Nos besoins en énergie sont énormes ; et comme je l'ai fait remarquer plus haut, il est possible que nous ne puissions plus compter sur les combustibles fossiles d'ici peu, ou que nous devions attendre longtemps leur remplacement massif par l'énergie solaire. L'énergie nucléaire, d'autre part, est à portée de la main, et il ne manque pas de voix autorisées pour nous expliquer qu'en prenant les précautions nécessaires, elle n'est *pas* plus dangereuse que les combustibles fossiles, qu'elle l'est même moins. (Dans le cas particulier de la contamination radioactive, on doit rappeler que le charbon contient de minuscules quantités d'impuretés radioactives, et que la combustion du charbon libère plus de radioactivité dans l'air que les réacteurs nucléaires ne le font – c'est du moins ce que certains prétendent.)

LES RÉACTEURS SURRÉGÉNÉRATEURS

Supposons que nous considérions la fission nucléaire comme une source d'énergie. Pour combien de temps pouvons-nous compter sur elle ? Pas pour très longtemps, si nous dépendons entièrement du matériau rare qu'est l'uranium 235 fissile. Mais heureusement, on a pu créer d'autres combustibles fissiles à partir de l'uranium 235 comme déclencheur de réaction.

Nous avons vu que le plutonium est l'un de ces combustibles dus à la main de l'homme. Supposons que nous construisions un petit réacteur à uranium enrichi sans y mettre de modérateur, de façon que des neutrons rapides arrivent dans une couverture d'uranium naturel entourant le tout. Si on s'arrange pour que peu de neutrons soient perdus, à partir de chaque fission d'un atome d'uranium 235 dans le cœur, on obtiendra plus d'un atome de plutonium créé dans la couverture.

Le premier *réacteur surrégénérateur* a été construit sous la direction du physicien américano-canadien Walter Henry Zinn à Arco, Idaho, en 1951. Il s'appelait EBR-1. Sa réalisation a prouvé la faisabilité des principes de la surrégénération, mais en outre il a produit de l'électricité. Il a été retiré du service pour obsolescence en 1964 (tant le progrès est rapide dans ce domaine).

La surrégénération pourrait multiplier les réserves en combustible à partir de l'uranium, puisque la totalité de l'isotope commun de l'uranium, l'uranium 238, devient un combustible potentiel.

L'élément thorium, constitué entièrement de thorium 232, est un autre combustible fissile potentiel. En absorbant des neutrons rapides, il devient l'isotope artificiel thorium 233, qui se désintègre ensuite en uranium 233. L'uranium 233, lui, est fissile sous l'effet des neutrons lents et il maintient une réaction en chaîne auto-entretenue. On peut donc ajouter du thorium aux réserves de combustible ; le thorium est environ cinq fois plus

abondant sur Terre que l'uranium. En fait, on a estimé que les cent mètres supérieurs de l'écorce terrestre contiennent en moyenne 5 000 tonnes d'uranium et de thorium au kilomètre carré. Bien sûr, tout ne peut être facilement extrait.

En bref, la quantité totale d'énergie disponible à partir des réserves d'uranium et de thorium du globe équivaudrait à environ vingt fois celle disponible à partir du charbon et du pétrole que nous n'avons pas encore extraits.

Pourtant, quand on leur parle de surrégénérateurs, les gens manifestent encore plus d'inquiétudes que lorsqu'il s'agit de réacteurs ordinaires. Le plutonium est beaucoup plus dangereux que l'uranium, et beaucoup maintiennent que le matériau le plus toxique du monde risque d'être produit en quantité massive, et que si une partie atteignait l'environnement, ce serait une catastrophe irréversible. Il y a aussi la crainte que le plutonium, destiné aux réacteurs à des fins pacifiques, puisse être détourné ou volé pour construire une bombe nucléaire (comme l'a fait l'Inde), qui pourrait être utilisée comme arme de chantage.

Ces frayeurs sont peut-être exagérées, mais elles sont en partie justifiées, et elles n'ont pas uniquement pour origines l'accident et le vol. Même si les réacteurs nucléaires fonctionnent sans accident, le danger subsistera. Pour en comprendre les raisons, examinons la radioactivité et le rayonnement énergétique auquel elle donne naissance.

LES DANGERS DES RADIATIONS

A vrai dire, la vie sur Terre a toujours été exposée à la radioactivité naturelle et aux rayons cosmiques. Cependant, la production de rayons X en laboratoire et la concentration de substances naturellement radioactives, comme le radium, qui existent ordinairement sous la forme de traces très diluées dans la croûte terrestre, ont considérablement augmenté le danger. Quelques-uns des premiers chercheurs à avoir travaillé sur les rayons X et le radium en ont reçu des doses léthales : Marie Curie et sa fille Irène Joliot-Curie sont mortes de leucémie due à leur exposition aux rayonnements, et l'on se souvient du cas des peintres de cadrans de montres dans les années vingt, qui sont morts pour avoir affiné avec leurs lèvres le bout de leurs pinceaux trempés dans le radium.

Le fait que le nombre de leucémies ait augmenté substantiellement dans les dernières décades peut être dû, partiellement, à l'augmentation de l'usage des rayons X à des fins diverses. L'incidence des leucémies dans le corps médical, qui est le plus exposé, est double par rapport à celle dans le public. Chez les radiologues, spécialistes médicaux de l'usage des rayons X, l'incidence est dix fois plus grande. Aussi a-t-on fait des essais pour substituer aux rayons X d'autres techniques, comme celles utilisant les ultrasons. L'avènement de la fission a beaucoup augmenté le danger. Que ce soit dans les bombes ou les réacteurs produisant de l'énergie, il y a libération de radioactivité sur une échelle qui pourrait rendre l'atmosphère tout entière, les océans, et tout ce que nous mangeons, buvons ou respirons, extrêmement dangereux pour la vie humaine. La fission a introduit une forme de pollution qui demandera aux hommes beaucoup d'intelligence pour la contrôler.

Quand les atomes d'uranium ou de plutonium se désintègrent, leurs *produits de fission* prennent des formes variées. Les fragments peuvent comprendre des isotopes du baryum, du technétium, entre autres possibilités. Pour tout dire, on a identifié quelque deux cents produits radioactifs de fission différents. Ils gênent l'industrie nucléaire, parce que certains absorbent fortement les neutrons et diminuent trop la réaction de fission. Pour cette raison, on doit extraire et purifier de temps en temps le combustible des réacteurs. En outre, les produits de fission sont tous dangereux pour la vie à des degrés divers, selon l'énergie et la nature du rayonnement. Il est plus dangereux par exemple d'avoir le corps exposé aux particules alpha qu'aux particules bêta. Le taux de désintégration joue aussi un rôle important : un nuclide qui se désintègre rapidement émettra plus de rayonnement à la seconde ou à l'heure qu'un autre qui se désintègre plus lentement.

Le taux de désintégration d'un nuclide radioactif n'a de sens que si un nombre important de ce nuclide est en jeu. Un noyau isolé peut se désintégrer à tout moment – tout de suite, dans un milliard d'années après ou n'importe quand entre ces deux moments – et il n'y a aucun moyen de prévoir quand il le fera. Chaque élément radioactif a cependant un taux moyen de désintégration, de sorte que si un grand nombre d'atomes est en jeu, il est possible de prédire avec une grande précision quelle proportion d'entre eux se sera désintégrée à un moment donné. Par exemple, disons que, d'après expérience, dans un échantillon donné d'un atome que nous appellerons X, les atomes se désintègrent à la vitesse de 1 sur 2 par an. A la fin de l'année, pour chaque groupe de 1 000 atomes X à l'origine dans l'échantillon, il en restera 500 ; au bout de deux ans 250, au bout de trois ans 125, et ainsi de suite. Le temps nécessaire pour que la moitié des atomes d'origine se désintègre s'appelle la *demi-vie* (expression introduite par Rutherford en 1904) de cet atome particulier ; par conséquent la demi-vie ou *période* de l'atome X est de un an. Chaque nuclide radioactif possède sa propre *demi-vie,* qui ne change jamais dans les conditions ordinaires. (La seule influence extérieure qui puisse apporter un changement est le bombardement du noyau par une particule ou la température extrêmement élevée régnant à l'intérieur des étoiles – en d'autres termes, un événement violent capable de modifier le noyau lui-même.)

La demi-vie de l'uranium 238 est de 4,5 milliards d'années. Il n'est donc pas surprenant qu'il reste de l'uranium 238 sur la Terre, en dépit de la désintégration des atomes d'uranium. Un calcul simple montre qu'il faudrait un temps plus long que six fois la demi-vie pour réduire une certaine quantité de nuclide radioactif à 1 % de sa valeur d'origine. Même dans 30 milliards d'années, il y aura encore un kilo d'uranium dans la croûte terrestre pour chaque tonne s'y trouvant actuellement.

Bien que les isotopes d'un élément soient pratiquement identiques chimiquement, ils peuvent différer énormément dans leurs propriétés nucléaires. L'uranium 235, par exemple, se désintègre six fois plus vite que l'uranium 238 ; sa demi-vie n'est que de 710 millions d'années. On peut donc en inférer qu'il y a des éternités, l'uranium était beaucoup plus riche en uranium 235 qu'il ne l'est aujourd'hui. Il y a six milliards d'années, par exemple, l'uranium 235 devait constituer environ 70 % de l'uranium naturel. L'humanité n'est pas pour autant en train de consommer les derniers restes de l'uranium 235. Même si nous avions été retardés d'un

million d'années dans notre découverte de la fission, la Terre aurait encore contenu 99,9 % de l'uranium 235 qu'elle contient actuellement.

Il est évident, en revanche, que tout nuclide ayant une demi-vie de moins de 100 millions d'années aurait diminué jusqu'à disparaître durant la longue existence de l'Univers. Nous ne pouvons donc trouver de nos jours que des traces de plutonium. En effet, l'isotope du plutonium ayant la plus longue demi-vie, le plutonium 244, a une demi-vie de seulement 70 millions d'années.

L'uranium, le thorium et d'autres éléments radioactifs de longue durée de vie, qui sont faiblement répandus dans les roches et le sol, produisent de petites quantités de radiations qui sont présentes dans l'air qui nous entoure. Nous-mêmes sommes légèrement radioactifs, car tous les tissus vivants contiennent des traces d'un isotope du potassium (le potassium 40), relativement rare et instable, qui a une demi-vie de un demi-milliard d'années. (Le potassium 40 produit, en se désintégrant, de l'argon 40, ce qui explique probablement que ce dernier est de loin le gaz inerte le plus répandu sur le globe. La mesure du rapport potassium-argon a été utilisée pour estimer l'âge des météorites.)

Il existe aussi un isotope radioactif du carbone, le carbone 14, qui ne devrait pas être normalement présent sur Terre puisque sa demi-vie n'est que de 5 770 ans. Cependant, du carbone 14 est continuellement formé dans notre atmosphère par le choc des particules des rayons cosmiques sur les atomes d'azote. Il en résulte que des traces de carbone 14 sont constamment présentes dans notre atmosphère, une partie s'incorporant au dioxyde de carbone. Comme il est présent dans le dioxyde de carbone, les plantes l'incorporent dans leurs tissus et il se répand à partir de là dans le monde animal, jusqu'à nous.

Le carbone 14, toujours présent dans le corps humain, y est en concentration bien plus faible que celle du potassium 40 ; mais le carbone 14, ayant une demi-vie bien plus courte, se désintègre beaucoup plus fréquemment. Le nombre total de désintégrations du carbone 14 peut être d'environ un sixième de celui des désintégrations du potassium 40. Cependant, il existe un certain pourcentage de carbone 14 dans les gènes humains, et quand il se désintègre, il peut en résulter de profondes modifications dans les cellules individuelles – modifications qui ne peuvent résulter des désintégrations du potassium 40.

On a pu avancer pour cette raison que le carbone 14 est le plus important des atomes radioactifs présents dans le corps humain. Ceci a été remarqué dès 1955 par le biochimiste américain d'origine russe Isaac Asimov.

Les divers nuclides radioactifs naturellement présents et les rayonnements énergétiques (comme les rayons cosmiques et les rayons gamma) sont à l'origine du *rayonnement de bruit de fond.* L'exposition constante au rayonnement naturel a probablement joué un rôle dans l'évolution en produisant des mutations, et elle est partiellement responsable de l'affection cancéreuse. Mais les organismes vivants ont vécu avec elle des milliards d'années. Le rayonnement nucléaire n'est devenu un danger sérieux qu'à notre époque, d'abord quand nous avons commencé à faire des expériences sur le radium, puis avec l'avènement de la fission et des réacteurs nucléaires. Au début du projet Manhattan, les physiciens avaient appris, à la suite d'expériences douloureuses, les dangers des rayonnements nucléaires. Les travailleurs du projet étaient donc astreints à des mesures de sécurité élaborées. Les produits « chauds » de fission et autres matériaux

radioactifs étaient placés derrière d'épais murs de blindage et on ne les regardait qu'à travers du verre au plomb. Chacun devait porter des bandes de film photographique ou d'autres dispositifs détecteurs pour contrôler la somme des expositions subies. On se livrait à un grand nombre d'expériences sur les animaux afin d'estimer l'*exposition maximale admissible.* (Les mammifères sont plus sensibles aux rayonnements que d'autres formes de vie ; parmi les mammifères, nous avons une résistance moyenne.)

En dépit de toutes les précautions prises, il s'est produit des accidents, et quelques physiciens nucléaires sont morts de *maladie des rayons* à la suite d'expositions à de fortes doses. Il subsiste pourtant des risques dans tout métier, même le plus sûr ; les travailleurs de l'industrie nucléaire sont réellement mieux protégés que la plupart des autres, grâce à une connaissance toujours plus complète des dangers et à l'attention portée pour les éviter. Mais dans un monde plein de réacteurs nucléaires qui rejetteront des milliers de tonnes de produits de fission, qu'en sera-t-il ? Comment se débarrassera-t-on de ces matériaux mortels ?

Une grande partie ne présente qu'une radioactivité à faible durée de vie qui s'atténue jusqu'à atteindre l'innocuité en quelques semaines ou quelques mois ; il suffit de les stocker durant ce temps puis de les jeter. Les nuclides ayant des demi-vies d'une à trente années sont plus dangereux. Ils ont une vie assez courte pour produire un rayonnement intense, et assez longue pour rester dangereux pendant plusieurs générations. Un nuclide ayant une demi-vie de trente ans mettra deux siècles à perdre 99 % de son activité.

L'UTILISATION DES PRODUITS DE FISSION

Les produits de fission peuvent être très utiles. Sources d'énergie, ils peuvent alimenter de petits dispositifs ou instruments. Les particules émises par l'isotope radioactif sont absorbées et leur énergie convertie en chaleur, celle-ci produisant à son tour de l'électricité dans des thermocouples. Les dispositifs produisant de l'électricité de cette façon sont des générateurs électriques à radio-isotopes, généralement appelés SNAPE (*s*ystèmes *n*ucléaires *a*uxiliaires pour la *p*roduction d'*é*nergie) ou encore, de façon plus évocatrice, *batteries atomiques.* Ils peuvent peser deux kilos à peine, fournir 60 watts et durer des années. On a utilisé des batteries SNAPE dans des satellites – dans *Transit 4A* et *Transit 4B,* par exemple, que les États-Unis ont mis en service en 1961 pour servir d'aides à la navigation.

L'isotope le plus couramment utilisé dans les batteries SNAPE est le strontium 90, dont nous reparlerons bientôt à propos d'autre chose. On utilise aussi des isotopes du plutonium et du curium.

Les astronautes qui sont allés sur la Lune ont placé à sa surface des générateurs nucléaires de ce type pour alimenter nombre d'expériences lunaires et d'équipements de radiotransmissions. Ces générateurs ont continué à fonctionner parfaitement durant des années.

Les produits de fission peuvent aussi trouver des utilisations potentielles en médecine (comme dans le traitement du cancer), dans la lutte anti-bactéries, la conservation des aliments et dans bien d'autres domaines de l'industrie, y compris la fabrication de produits chimiques. Par exemple, une société américaine, Hercules Powder Company, a mis au point un réacteur pour utiliser son rayonnement dans la production de l'antigel éthylène glycol.

Pourtant, quand on a bien fait le tour du problème, on voit que tous ces usages concevables ne consommeront qu'une petite partie des énormes quantités de produits de fission que les centrales nucléaires rejetteront. Cela représente une des grandes difficultés de l'électronucléaire en général. On estime que chaque tranche de 200 000 kilowatts de production d'électricité par le nucléaire entraînera la production de 700 grammes de produits de fission par jour. Qu'en faire ? Les États-Unis ont déjà stocké des millions de litres de liquides radioactifs dans le sous-sol ; on a estimé que, vers l'an 2000, il faudra se débarrasser de quelque deux millions de litres de liquide radioactif chaque jour ! Les États-Unis et la Grande-Bretagne ont jeté en mer des récipients en béton contenant des produits de fission. On a proposé de jeter les déchets radioactifs dans les abysses océaniques, de les stocker dans des mines de sel désaffectées, de les enfermer dans du verre fondu et d'enterrer les matériaux solides. Mais il subsiste toujours l'obsession de voir d'une façon ou d'une autre la radioactivité s'échapper et contaminer le sol ou les mers. C'est avec horreur que l'on songe à la possibilité de voir un jour un navire à propulsion nucléaire sombrer et répandre ses produits de fission dans l'Océan. Le naufrage du sous-marin nucléaire américain *Thresher* dans l'Atlantique nord le 10 avril 1963 a particulièrement alimenté ce genre de peur, bien que dans ce cas précis une telle contamination ne se soit pas produite.

LES RETOMBÉES RADIOACTIVES

Si la pollution radioactive par l'énergie nucléaire pacifique constitue un danger potentiel, au moins on s'efforce de la contrôler par tous les moyens possibles, et probablement avec succès. Mais il y a une pollution qui s'est déjà répandue sur le globe terrestre et qui d'ailleurs, dans une guerre nucléaire, pourrait être délibérément créée. Il s'agit des retombées des bombes atomiques.

Les retombées sont le lot de toutes les bombes nucléaires, même celles qui ne sont pas mises à feu durant un conflit. Comme les retombées sont transportées autour du globe par les vents et ramenées à terre par la pluie, il est presque impossible pour une nation de faire exploser une bombe nucléaire dans l'atmosphère sans qu'elle soit détectée. Dans l'éventualité d'une guerre nucléaire, les retombées à longue échéance peuvent produire plus d'accidents et causer plus de dommages aux êtres vivants de la planète que la chaleur et le souffle de la bombe elle-même n'en feraient subir aux contrées attaquées.

On divise les retombées en trois catégories, les retombées *locales, troposphériques* et *stratosphériques.* Les retombées locales résultent des explosions au sol, dans lesquelles des isotopes radioactifs sont adsorbés sur des particules du sol et rapidement répandus à 160 kilomètres de l'explosion. Le souffle des bombes à fission d'une taille de l'ordre de la kilotonne envoie des produits de fission dans la troposphère. Ceux-ci y séjournent environ un mois tout en étant pendant ce temps transportés par les vents à quelques milliers de kilomètres vers l'est.

L'énorme quantité de produits de fission libérés par les superbombes thermonucléaires est expédiée jusque dans la stratosphère. Ces retombées stratosphériques y séjournent un an ou plus, se répartissent sur un hémisphère entier et retombent finalement aussi bien sur l'attaquant que sur l'attaqué.

L'intensité des retombées de la première superbombe, qui a explosé le 1er mars 1964 dans le Pacifique, a pris les scientifiques par surprise. Ils ne s'attendaient pas à ce que les retombées d'une bombe à fusion soient si « sales ». Trois mille kilomètres carrés furent gravement contaminés – une surface presque de la taille de l'État du Massachusetts. Mais tout s'est éclairci quand on a appris que le cœur en fusion avait été complété par une couverture d'uranium 238 dont les neutrons avaient déclenché la fission. Ceci avait non seulement multiplié la force de l'explosion, mais avait aussi donné naissance à un nuage de produits de fission bien plus grand que celui d'une simple bombe à fission du type de celle d'Hiroshima. Les retombées des essais de bombes ont élevé d'une faible quantité seulement la radioactivité du rayonnement de bruit de fond de la Terre. Mais une faible augmentation du niveau naturel de rayonnement peut augmenter les risques de cancer, entraîner des accidents génétiques et raccourcir légèrement l'espérance de vie humaine. Les estimations les plus modestes de ces risques reconnaissent que, en augmentant le taux de mutations (voir à ce sujet le chapitre 13), les retombées constituent un danger non négligeable pour les générations futures.

L'un des produits de fission est particulièrement dangereux pour la vie humaine. C'est le strontium 90 (demi-vie de vingt-huit ans), l'isotope si utile aux générateurs SNAPE. Le strontium 90 tombant sur le sol et dans l'eau est assimilé par les plantes et parvient peu après dans le corps des animaux (incluant l'homme) qui se nourrissent directement ou indirectement des végétaux. Le strontium est particulièrement dangereux dans la mesure où, à cause de sa parenté chimique avec le calcium, il parvient dans les os où il se fixe pour une longue durée. Les minéraux, dans les os, ont un faible taux de renouvellement, c'est-à-dire qu'ils ne sont pas remplacés aussi rapidement que les substances des autres tissus animaux. C'est pour cette raison que le strontium 90, une fois absorbé, peut rester dans le corps pendant une grande partie de la vie (figure 10.5).

Le strontium 90 est une substance toute nouvelle dans notre environnement ; il n'existait sur Terre en aucune quantité détectable jusqu'à ce que les scientifiques réalisent la fission de l'atome d'uranium. Mais aujourd'hui, en moins d'une génération, un peu de strontium 90 s'est fixé dans les os de chaque être humain du globe, et d'ailleurs de tous les vertébrés. Il en flotte encore des quantités considérables dans notre stratosphère, qui viendront tôt ou tard élever la concentration existant déjà dans nos os.

Le strontium 90 se mesure en *unités de strontium* (US). Une US vaut une micromicrocurie de strontium 90 par gramme de calcium dans le corps. Une *curie* est une unité de rayonnement (ainsi nommée en l'honneur des Curie, bien sûr) qui était, à l'origine, équivalente au rayonnement produit par un gramme de radium en équilibre avec son produit de désintégration, le radon, mais actuellement la curie est définie comme le rayonnement dû à 37 milliards de désintégrations par seconde. Une micromicrocurie est un mille milliardième de curie, soit 2,22 désintégrations par minute. Une unité de strontium représente donc 2,22 désintégrations par minute par gramme de calcium présent dans le corps. La concentration de strontium 90 dans le squelette humain varie énormément selon le lieu et la personne. On a trouvé chez certains sujets des concentrations de soixante-cinq fois la valeur moyenne. Chez les enfants, la valeur moyenne de concentration observée vaut au moins quatre fois celle des adultes, en raison de la rotation plus rapide des éléments dans leurs os en croissance.

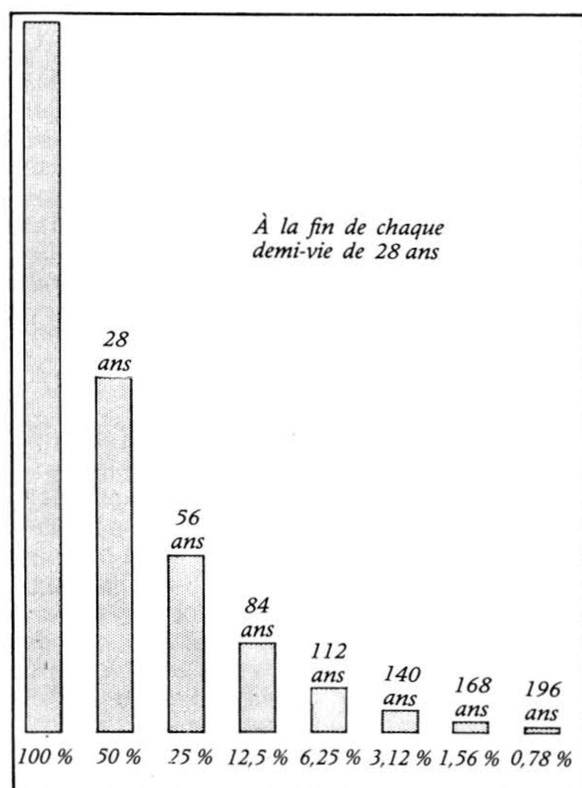


Figure 10.5. Désintégration du strontium 90 sur environ deux cents ans.

Les estimations des moyennes varient elles-mêmes beaucoup, parce qu'elles sont basées principalement sur des estimations de la quantité de strontium 90 trouvée dans l'alimentation. (Soit dit en passant, le lait n'est pas une nourriture particulièrement dangereuse à cet égard bien que le calcium provenant des végétaux ait plus de strontium 90 associé. Ce que l'on peut appeler le *système de filtrage* de la vache élimine une partie du strontium qu'elle trouve dans ses aliments végétaux.) L'estimation de la concentration moyenne de strontium 90 dans les os des habitants des États-Unis en 1959, avant que soient arrêtées les explosions nucléaires atmosphériques, allait de moins d'une unité de strontium jusqu'à plus de cinq unités. (Le *maximum admissible* établi par la Commission internationale de radioprotection était de 67 US.) Mais les moyennes ne signifient pas grand-chose, d'autant moins que le strontium 90 peut se concentrer dans des *points chauds* dans les os et y atteindre un niveau assez élevé de concentration pour déclencher une leucémie ou un cancer. L'importance des effets des rayonnements a, entre autres, entraîné l'adoption de plusieurs types d'unités pour les mesurer. L'une d'elles, le *roentgen,* ainsi nommée en l'honneur de l'inventeur des rayons X, est basée sur le nombre d'ions

produits par les rayons X ou gamma mesurés. Plus récemment, on a introduit le *rad* (raccourci de « radiation »). Il représente l'absorption de 100 ergs par gramme d'un type quelconque de rayonnement.

La nature du rayonnement est d'importance. Un rad de particules de masse élevée est bien plus efficace pour induire des changements chimiques dans les tissus qu'un rad de particules légères ; l'énergie sous forme de particules alpha est bien plus dangereuse que sous forme d'électrons.

Chimiquement, les dommages dus au rayonnement sont principalement créés par la cassure des molécules d'eau (qui représentent la majeure partie de la masse des tissus vivants) en fragments hautement actifs (*radicaux libres*) qui, à leur tour, réagissent avec les molécules complexes des tissus. Les atteintes à la moelle des os, interférant avec la production des cellules sanguines, constituent l'une des plus sérieuses manifestations de la maladie des rayons et, si elles sont assez importantes, elles sont irréversibles et conduisent à la mort.

De nombreux et éminents scientifiques pensent fermement que les retombées des essais des bombes nucléaires sont un péril pour l'espèce humaine. Le chimiste américain Linus Pauling a avancé que les retombées d'une simple superbombe peuvent entraîner jusqu'à 100 000 morts par leucémie et autres maladies sur tout le globe, et il a fait remarquer que le carbone 14 radioactif, produit par les neutrons d'une explosion nucléaire, constitue un sérieux danger génétique. Il a, pour cette raison, été extrêmement actif pour demander la cessation des essais de bombes nucléaires ; il a soutenu tous les mouvements ayant pour buts de diminuer le risque de guerre et d'encourager le désarmement. Mais d'autre part, quelques scientifiques, dont le physicien américain d'origine hongroise Edward Teller, ont cherché à minimiser les risques des retombées. La sympathie de l'opinion publique mondiale s'est plutôt portée du côté de Pauling ; en témoigne peut-être le fait qu'il a reçu en 1963 le prix Nobel de la paix.

A l'automne 1958, les États-Unis, l'Union Soviétique et la Grande-Bretagne ont suspendu les essais de bombes nucléaires d'un commun accord. (Cela n'a cependant pas empêché la France de faire exploser sa première bombe thermonucléaire dans l'atmosphère au printemps 1960.) Pendant trois ans, la conjoncture a paru meilleure ; la concentration de strontium 90 a atteint un maximum et s'est stabilisée vers 1960 à un point bien en dessous de celui estimé être le maximum compatible avec la sécurité. Même ainsi, quelque 25 millions de curies de strontium 90 et de césium 137 (un autre produit de fission dangereux) ont été envoyées dans l'atmosphère durant les trente années d'essais nucléaires pendant lesquelles ont explosé quelque 150 bombes de modèles variés. Deux seulement de ces bombes ont explosé en période de conflit, mais les effets en ont été effroyables.

En 1961, sans avertissement, l'Union soviétique a mis un terme au moratoire et repris les essais. Comme l'Union Soviétique faisait exploser des bombes thermonucléaires d'une puissance sans précédent, les États-Unis se sont sentis contraints à recommencer les essais. L'opinion mondiale, sensibilisée par le soulagement apporté par le moratoire, a réagi avec indignation.

Le 10 octobre 1963, cependant, les trois principales puissances nucléaires ont signé un *traité* (et pas un simple accord) pour l'interdiction partielle

des essais nucléaires, selon lequel les explosions de bombes nucléaires dans l'atmosphère, dans l'espace ou dans la mer étaient interdites. Seules étaient autorisées les explosions souterraines puisqu'elles ne produisent pas de retombées. Ce traité est la démarche la plus encourageante pour la survie de l'espèce que l'humanité ait accomplie depuis le début de l'ère nucléaire.

La fusion nucléaire contrôlée

Depuis plus de trente ans, les physiciens nucléaires ont eu à l'esprit un rêve encore plus stimulant que celui d'utiliser la fission à des usages constructifs ; c'était celui de maîtriser l'énergie de la fusion. La fusion, après tout, est ce qui fait marcher notre monde : les réactions de fusion à l'intérieur du Soleil sont à l'origine de toutes les formes d'énergie et de la vie elle-même. Si nous pouvions reproduire et contrôler de telles réactions sur la Terre, tous nos problèmes d'énergie seraient résolus. Notre réserve de combustible serait aussi grande que l'Océan, puisque le combustible serait l'hydrogène.

Curieusement, ce ne serait pas le premier usage de l'hydrogène en tant que combustible. Peu de temps après qu'on eut découvert l'hydrogène et étudié ses propriétés, il fut utilisé comme combustible chimique. Le scientifique américain Robert Hare mit au point en 1801 une torche oxhydrique, et la flamme chaude fournie par la combustion de l'hydrogène dans l'oxygène a toujours été utilisée depuis dans l'industrie.

L'hydrogène liquide a aussi été utilisé comme un combustible d'une immense importance dans les fusées, et on a envisagé de l'utiliser en tant que combustible particulièrement propre pour la génération d'électricité et la propulsion des automobiles et autres véhicules. (Dans ces derniers exemples, le problème de sa facilité d'explosion dans l'air demeure.) C'est, cependant, en tant que combustible pour la fusion nucléaire que son avenir paraît le plus brillant.

L'énergie de la fusion serait immensément plus commode que celle de la fission. A poids égal, un réacteur à fusion fournirait cinq à dix fois plus d'énergie qu'un réacteur à fission. Un kilogramme d'hydrogène pourrait produire par la fusion 70 millions de kilowatt-heures d'énergie. En outre, la fusion dépendrait d'isotopes de l'hydrogène que l'on peut facilement tirer de l'Océan en grande quantité, tandis que la fission nécessite l'extraction d'uranium et de thorium du sous-sol – une tâche comparativement beaucoup plus difficile. En outre, la fusion ne produit que des neutrons et de l'hydrogène 3, qui sont beaucoup moins dangereux à contrôler que les produits de fission. Finalement, et c'est peut-être le plus important, une réaction de fusion, en cas d'un quelconque incident de fonctionnement, s'arrêterait tout simplement, tandis qu'une réaction de fission peut échapper aux contrôles (*l'excursion nucléaire*), produire une *fonte* de son uranium (bien que ce ne soit encore jamais arrivé), et répandre des produits radioactifs.

Si on arrivait à la fusion nucléaire contrôlée, elle pourrait, étant donné la disponibilité en combustible et la richesse de l'énergie qu'elle produirait, constituer une source d'énergie utile qui durerait des milliards d'années

III. La technique

Planche III.1. Outils en pierre des hommes préhistoriques. Les plus anciens, datant de la période du Miocène, sont *en bas à gauche* ; les plus récents *en bas à droite.* (Nég. n° 411257, photo J. Kirschner ; avec l'aimable autorisation de la bibliothèque du Musée national américain d'histoire naturelle.)

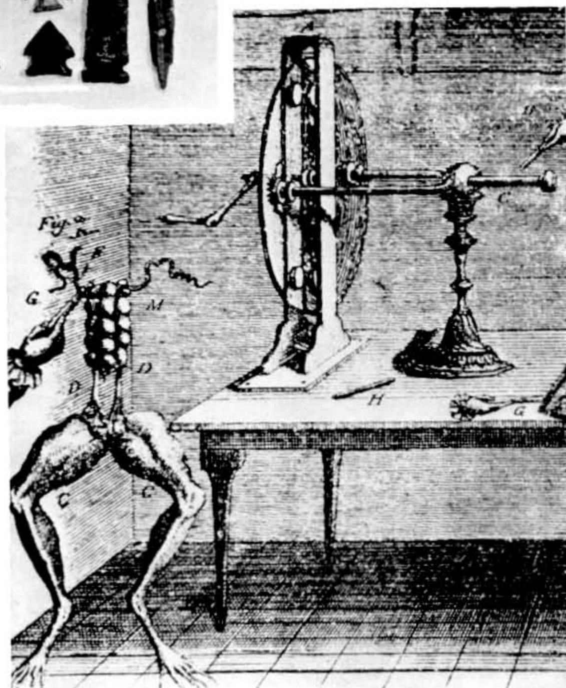


Planche III.2. L'expérience de Galvani, qui conduisit à la découverte des courants électriques. L'électricité produite par sa machine électrostatique faisait se contracter la patte de la grenouille ; il trouva qu'en touchant le nerf avec deux métaux différents, on faisait aussi se contracter la patte de la grenouille. (Avec l'aimable autorisation des archives Bettmann.)

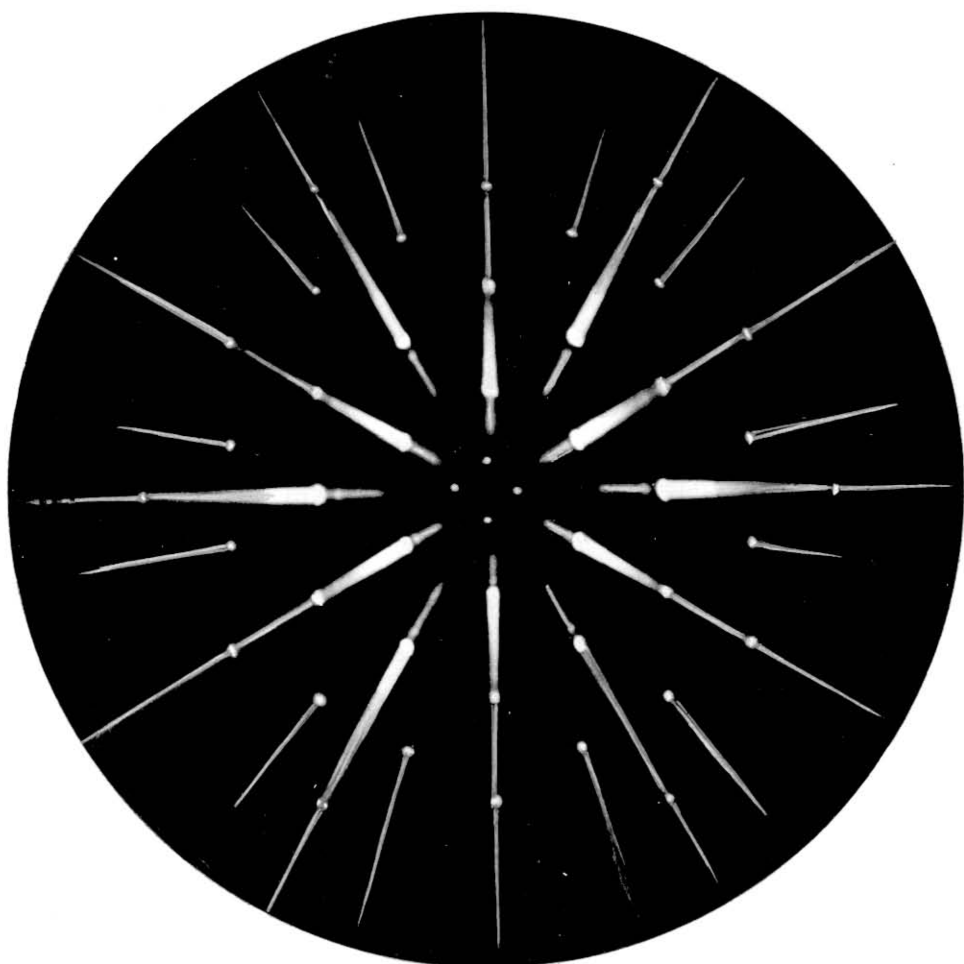


Planche III.3. Un cristal de glace isolé photographié par diffraction des rayons X, montrant la symétrie et l'équilibre des forces physiques maintenant la structure. (Franklin Branley, éd. *Scientists/Choice*, New York Basic Books, s.d.)

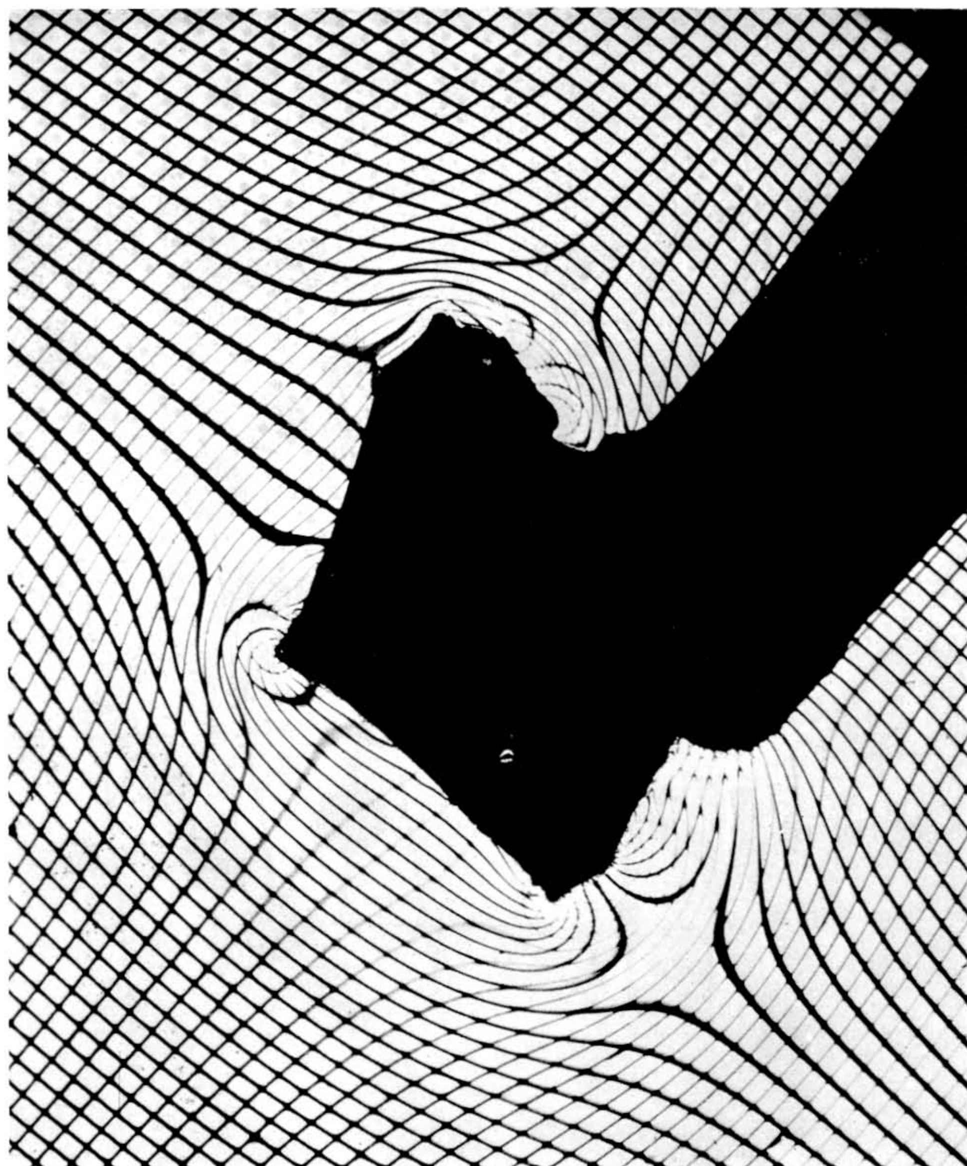


Planche III.4. Le champ électrique autour d'un cristal chargé est photographié au microscope électronique au moyen d'une technique d'ombre. La méthode utilise une grille métallique fine ; la distorsion du maillage, due à la déflection des électrons, montre la forme et l'intensité du champ électrique. (Avec l'aimable autorisation du Bureau national américain des poids et mesures.)

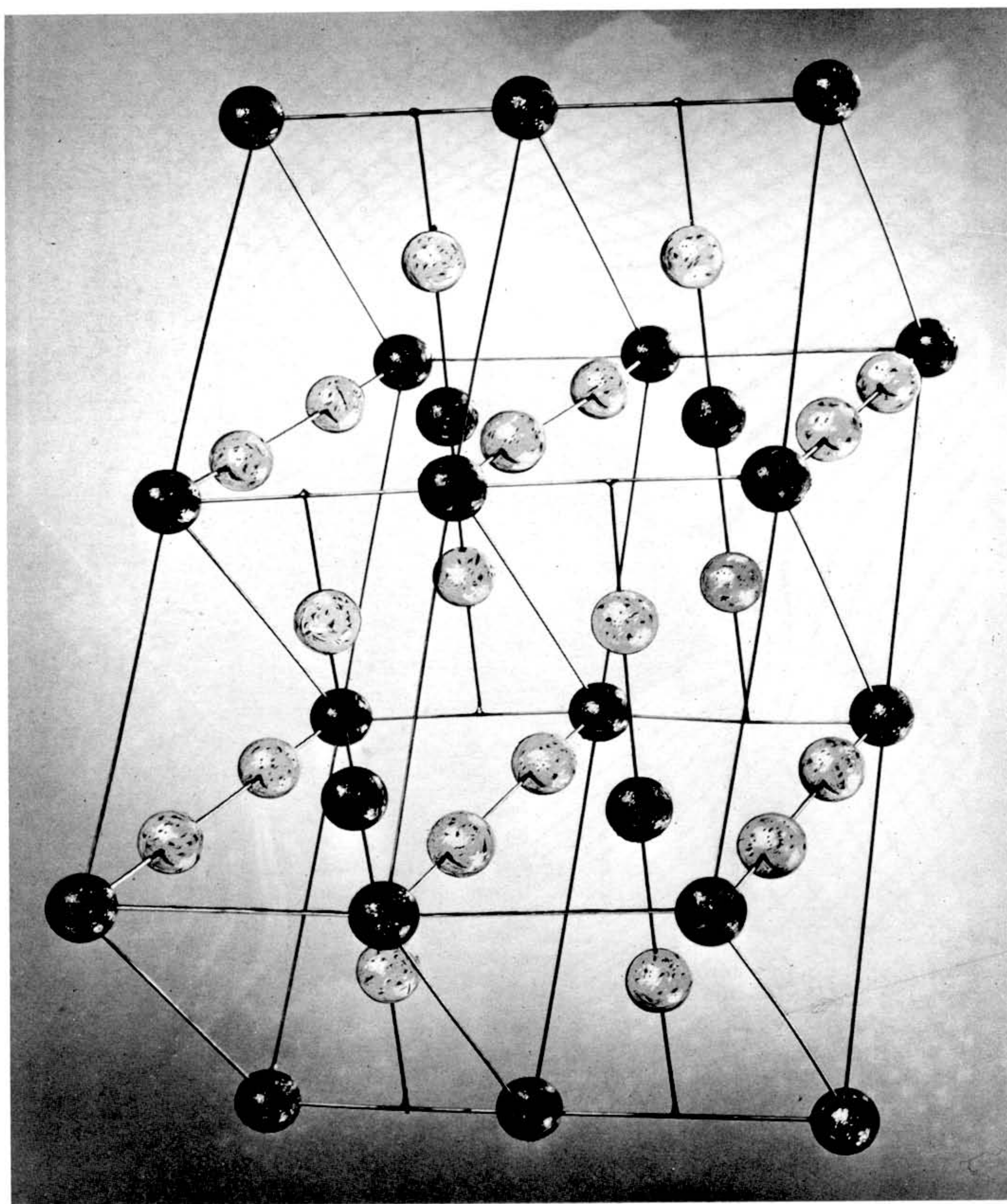


Planche III.5. Modèle moléculaire de l'oxyde de titane, qui peut servir de semi-conducteur, sous sa forme cristalline. La suppression de l'un des atomes d'oxygène (boules claires) rendra le matériau semi-conducteur. (Avec l'aimable autorisation du Bureau national américain des poids et mesures.)

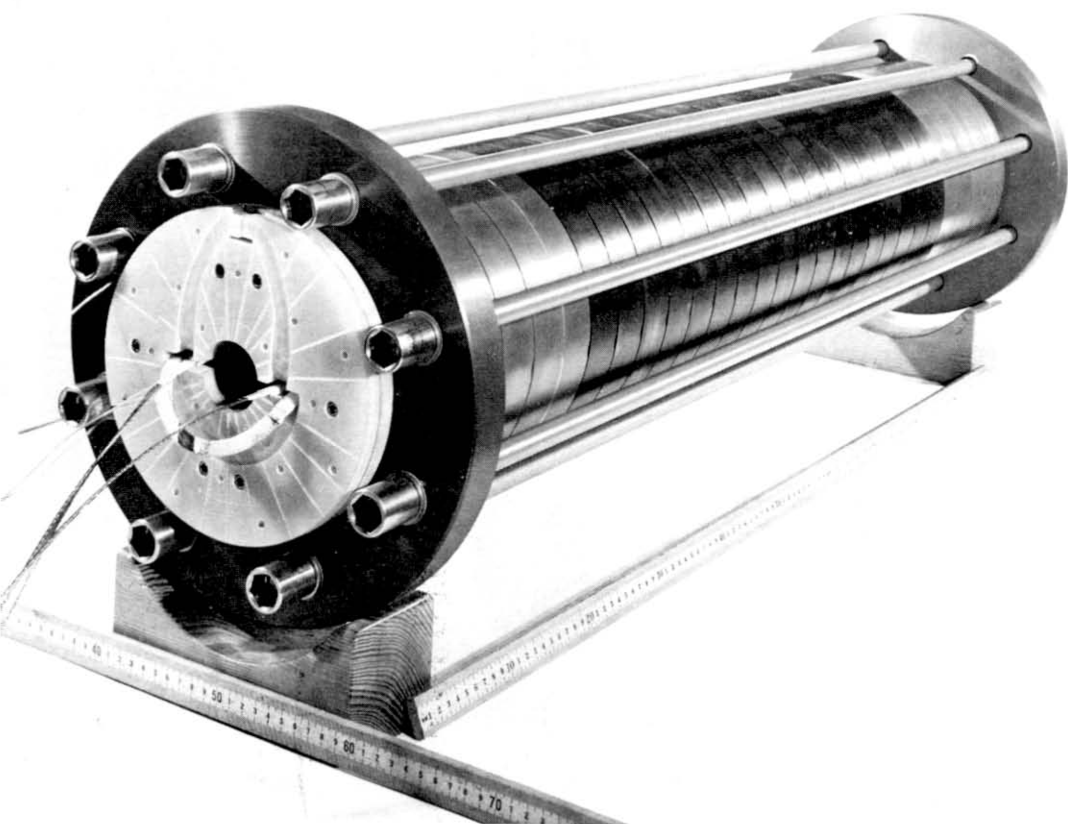


Planche III.6. Dans un cyclotron, on utilise des aimants pour courber un faisceau de particules chargées électriquement le long d'une trajectoire circulaire. Avec l'augmentation des besoins scientifiques en faisceaux d'énergie toujours plus élevée, ces accélérateurs et leurs aimants ont vu leur taille considérablement augmenter. Ici, un modèle d'aimant supraconducteur mis au point par Clyde Taylor et ses collaborateurs au laboratoire Lawrence Berkeley. (Avec l'aimable autorisation du laboratoire Lawrence Berkeley, université de Californie, Berkeley.)

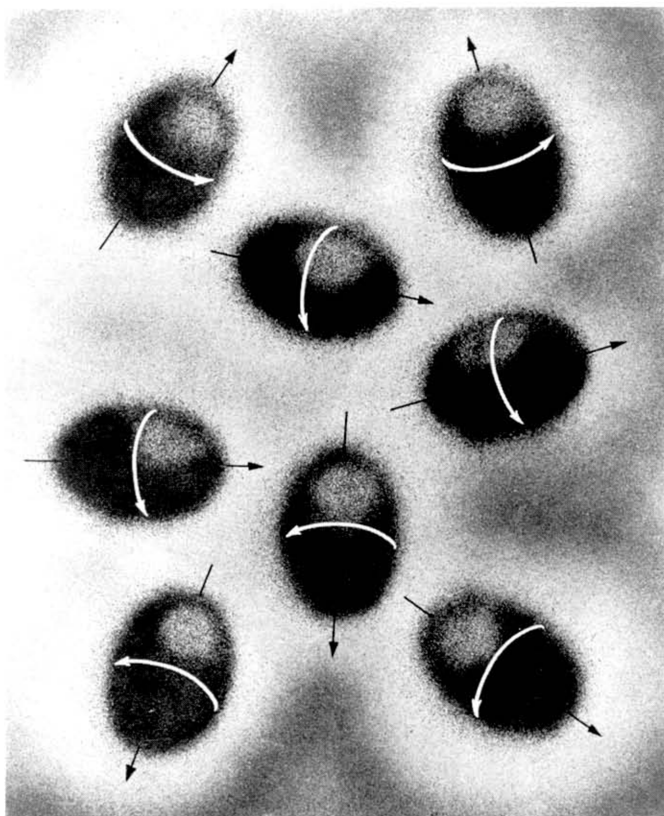


Planche III.7. On a représenté sur ce schéma des protons dont les axes de rotation sur eux-mêmes sont orientés dans des directions aléatoires. La flèche noire représente la direction du spin du proton. (Avec l'aimable autorisation du Bureau national américain des poids et mesures.)

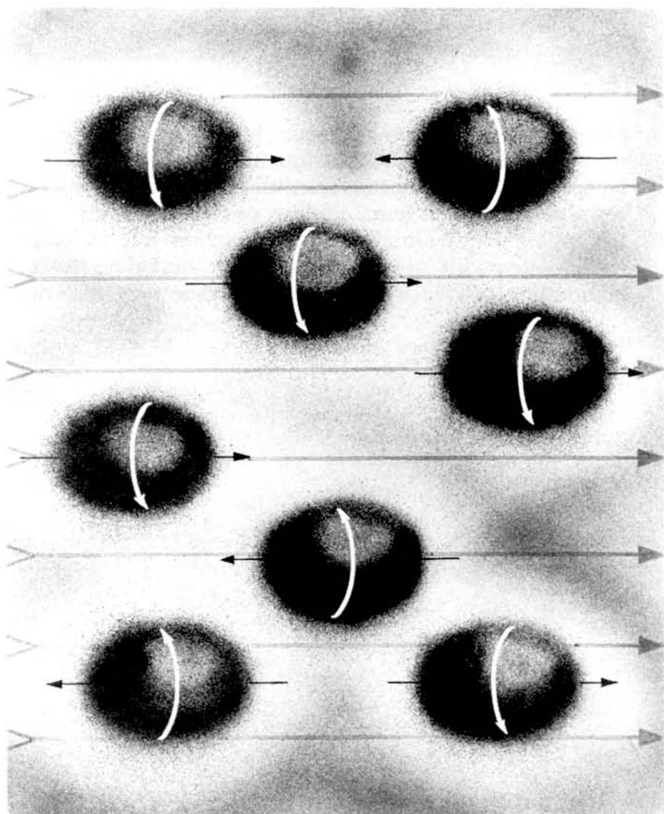


Planche III.8. Protons aux spins alignés par un champ magnétique constant. Ceux orientés dans la direction opposée à la normale (*flèches noires pointées vers la gauche*) sont dans l'état de spin excité. (Avec l'aimable autorisation du Bureau national américain des poids et mesures.)

Planche III.9. Traces d'électrons et de positrons créés dans une chambre à bulles par des rayons gamma de haute énergie. Les cercles sont dus à des électrons soumis à un champ magnétique. (Avec l'aimable autorisation de l'université de Californie, Berkeley.)

Planche III.10. Fission d'un atome d'uranium. Le segment blanc au milieu de cette photographie représente les trajectoires de deux atomes s'écartant l'un de l'autre à partir du point central où l'atome d'uranium s'est scindé en deux. La plaque photographique a été trempée dans un composé d'uranium et bombardée avec des neutrons qui ont causé la fission révélée sur cette photo. Les autres tâches blanches sont des grains d'argent qui se sont développés de façon aléatoire. La photo a été faite aux laboratoires de recherche Eastman Kodak. (Avec l'aimable autorisation de United Press International.)



Planche III.11. Visualisation de la radio-activité. Dans le seau, il y a du tantale, rendu radioactif dans le réacteur de Brookhaven ; ce matériau, qui émet de la lumière, est recouvert, par précaution, de plusieurs dizaines de centimètres d'eau. Le tantale radioactif sera placé dans le tuyau que l'on aperçoit et transféré dans un grand récipient en plomb pour être utilisé comme source de radioactivité de 1 000 curies à des fins industrielles. (Avec l'aimable autorisation du laboratoire national de Brookhaven, New York.)

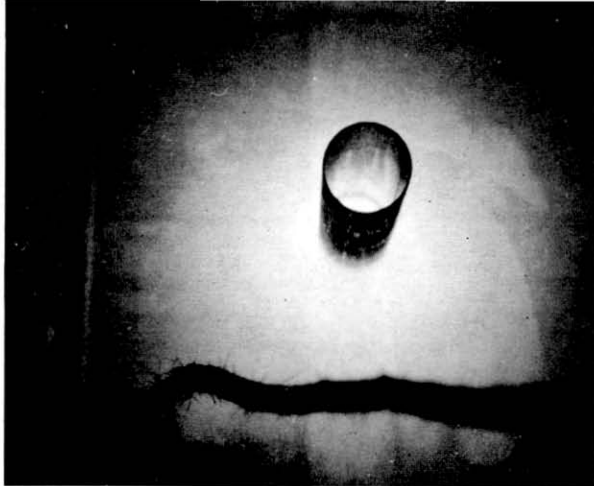
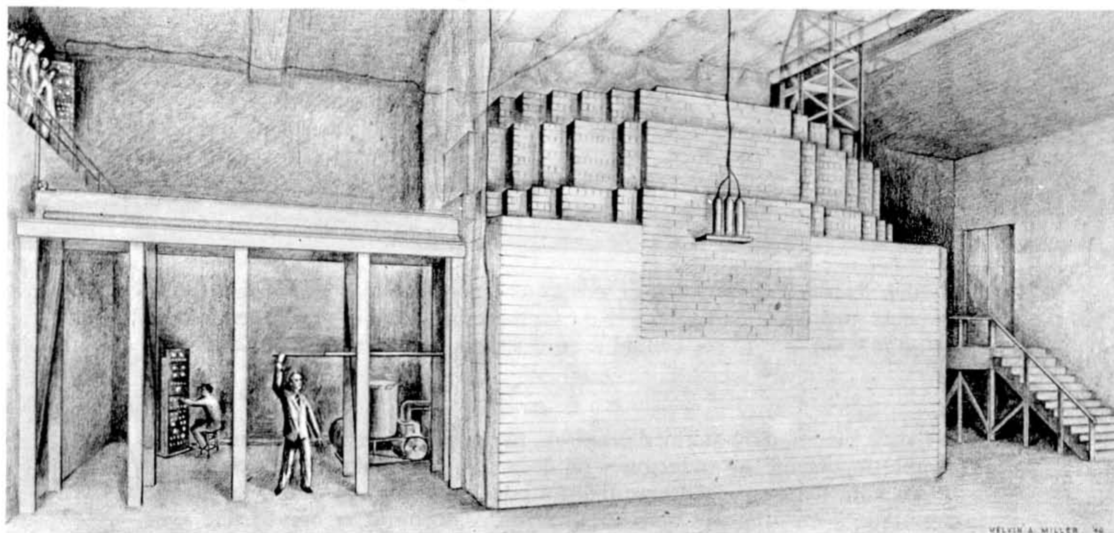


Planche III.12. Dessin du premier réacteur en chaîne, bâti sous le stade de football de Chicago. (Avec l'aimable autorisation du laboratoire national d'Argonne, Illinois.)



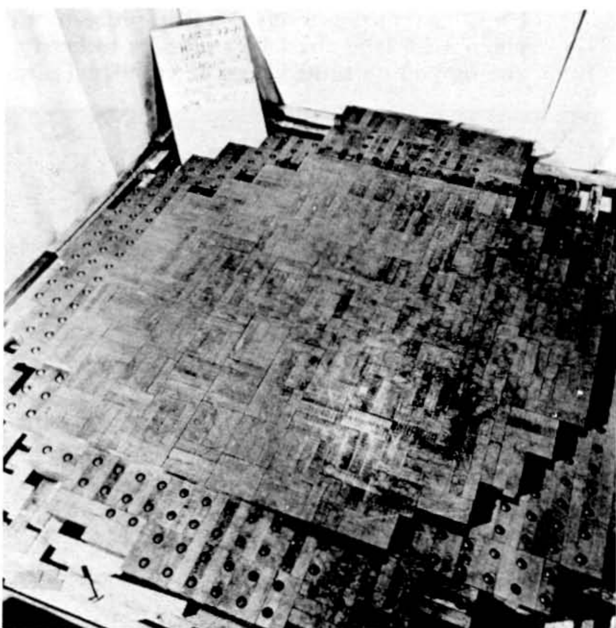


Planche III.13. Le réacteur de Chicago durant sa construction. Cette photo est l'une des rares prises durant la construction du réacteur. Les barres dans les trous sont de l'uranium, et on est en train de poser la dix-neuvième couche du réacteur, qui est constituée de blocs de graphite solide. (Avec l'aimable autorisation du laboratoire national d'Argonne, Illinois.)



Planche III. 14. Le nuage en forme de champignon de la bombe atomique lancée par les États-Unis sur Hiroshima, au Japon, le 6 août 1945. Cette photo a été prise par Seizo Yamada, un lycéen à l'époque. (Avec l'aimable autorisation du photographe.)

Planche III.15. Une tour de refroidissement du réacteur nucléaire de Three Mile Island, en Pennsylvanie, aux États-Unis. (Photo de Sylvia Plachy, avec son aimable autorisation.)

Planche III.16. La vie et la mort d'un plasma en pincement. Cette série de photos montre la brève histoire d'une striction de plasma dans le champ magnétique du Perhapsatron. Chaque photographie donne deux vues du plasma, l'une de côté et l'une de dessous grâce à un miroir. La striction se fragmente en un millionième de seconde ; le nombre sur chaque photo exprime des microsecondes. (Avec l'aimable autorisation du laboratoire scientifique de Los Alamos, Nouveau-Mexique.)

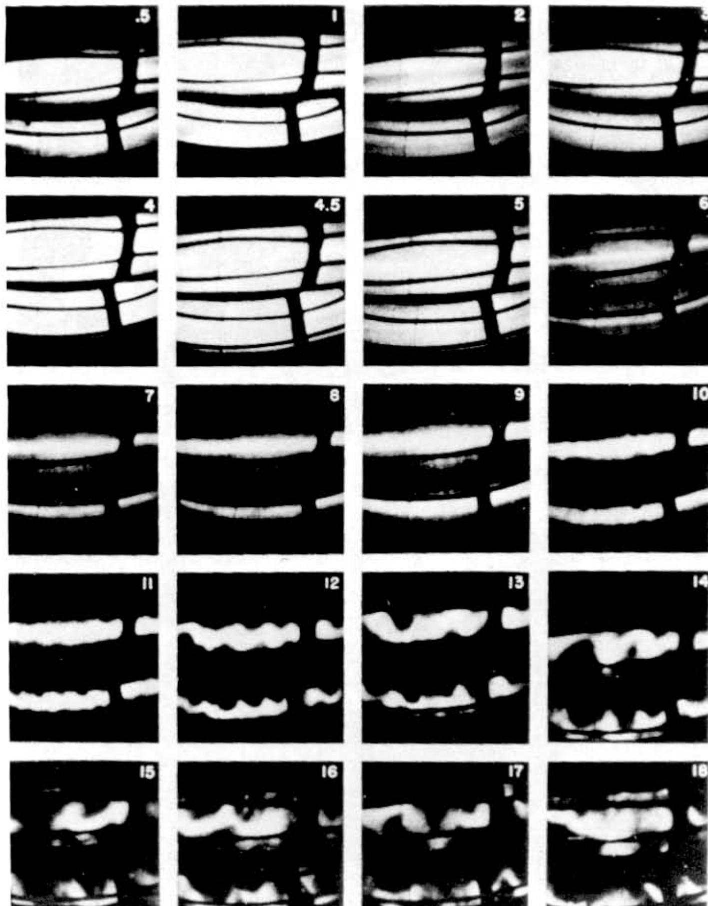
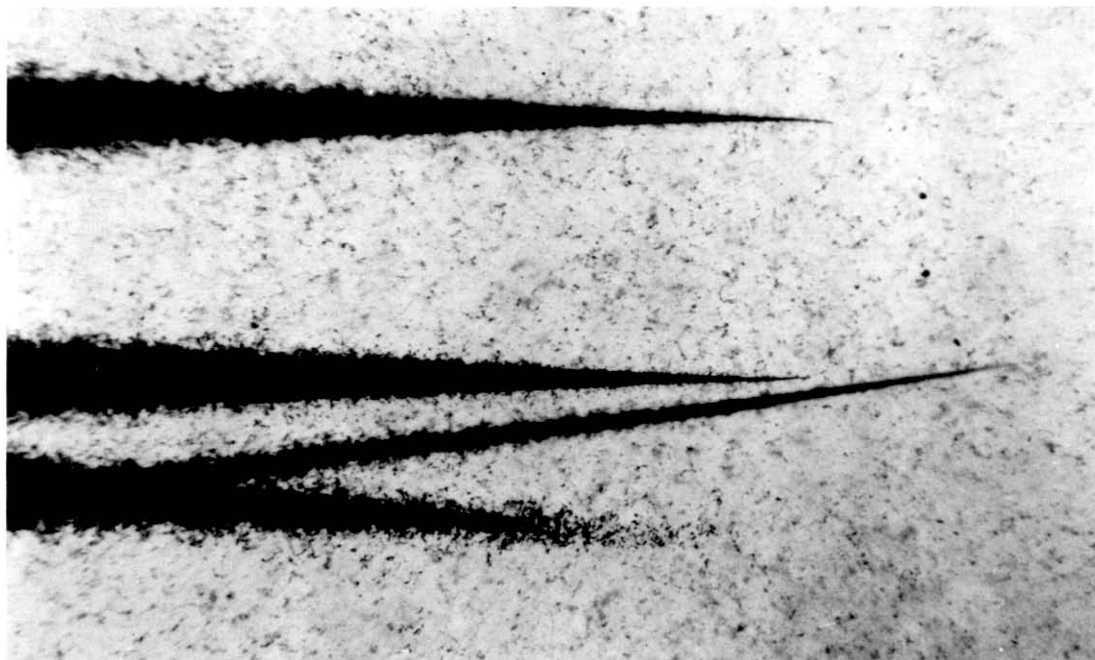


Planche III.17. Les traînées sombres sont les traces laissées par les premiers atomes d'uranium accélérés près de la vitesse de la lumière. On voit ici les derniers demi-millimètres de trois trajectoires, alors que les noyaux arrivent au repos dans une émulsion photographique spéciale. La trace du bas montre un noyau se scindant en deux noyaux plus légers. Cette expérience a été réalisée au Bevalac, le seul accélérateur de particules au monde qui puisse fournir des ions aussi lourds que l'uranium à des énergies relativistes. L'accélérateur est situé au laboratoire Lawrence Berkeley de l'université de Californie. (Avec l'aimable autorisation du laboratoire Lawrence Berkeley, université de Californie, Berkeley.)



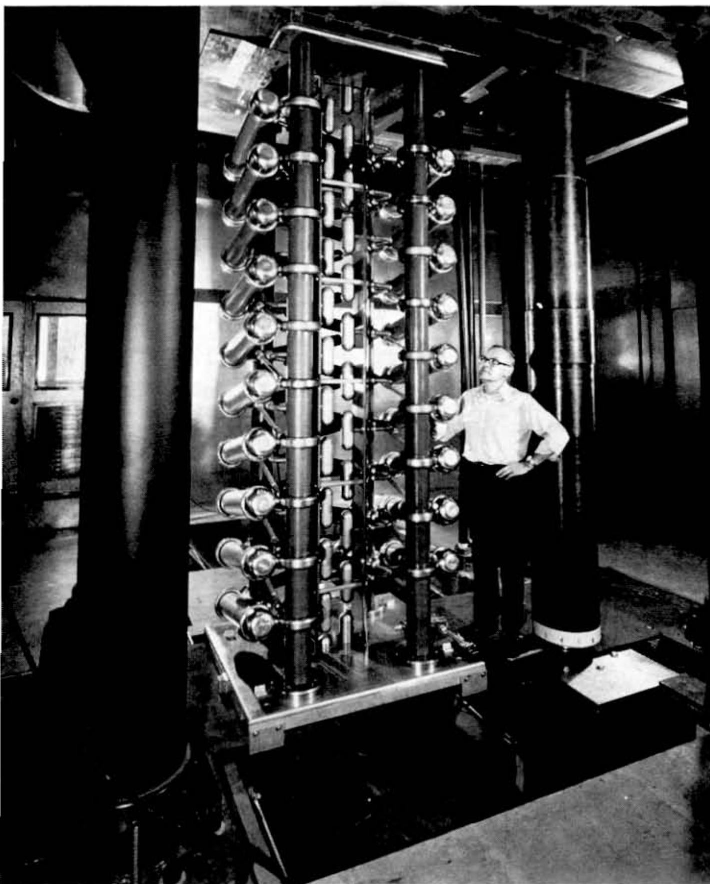


Planche III.18. Le banc de redresseurs qui permet l'alimentation à haute tension d'un nouveau système injecteur d'ions au Super HILAC du laboratoire Lawrence Berkeley. L'injecteur, appelé Abel, augmente les possibilités de l'accélérateur, ce qui lui permet de fournir des faisceaux d'ions lourds comme l'uranium. (Avec l'aimable autorisation du laboratoire Lawrence Berkeley, université de Californie, Berkeley.)

Planche III.19. La coquille d'aluminium placée au sommet du banc de redresseurs contient un aimant et une source d'ions, dans laquelle un champ électrique arrache les électrons des atomes. Les atomes ainsi traités (appelés ions) ont alors une charge positive et peuvent être accélérés par le champ électrique dans les colonnes d'accélération (visibles dans leur enceinte en plexiglas). Le faisceau d'ions subit une accélération supplémentaire, dans un nouvel accélérateur Wideroe, par le super HILAC, avant d'être envoyé dans le Bevatron. Quand le super HILAC et le Bevatron sont ainsi utilisés en tandem, on appelle cette combinaison le Bevalac. (Lab. L. Berkeley.)

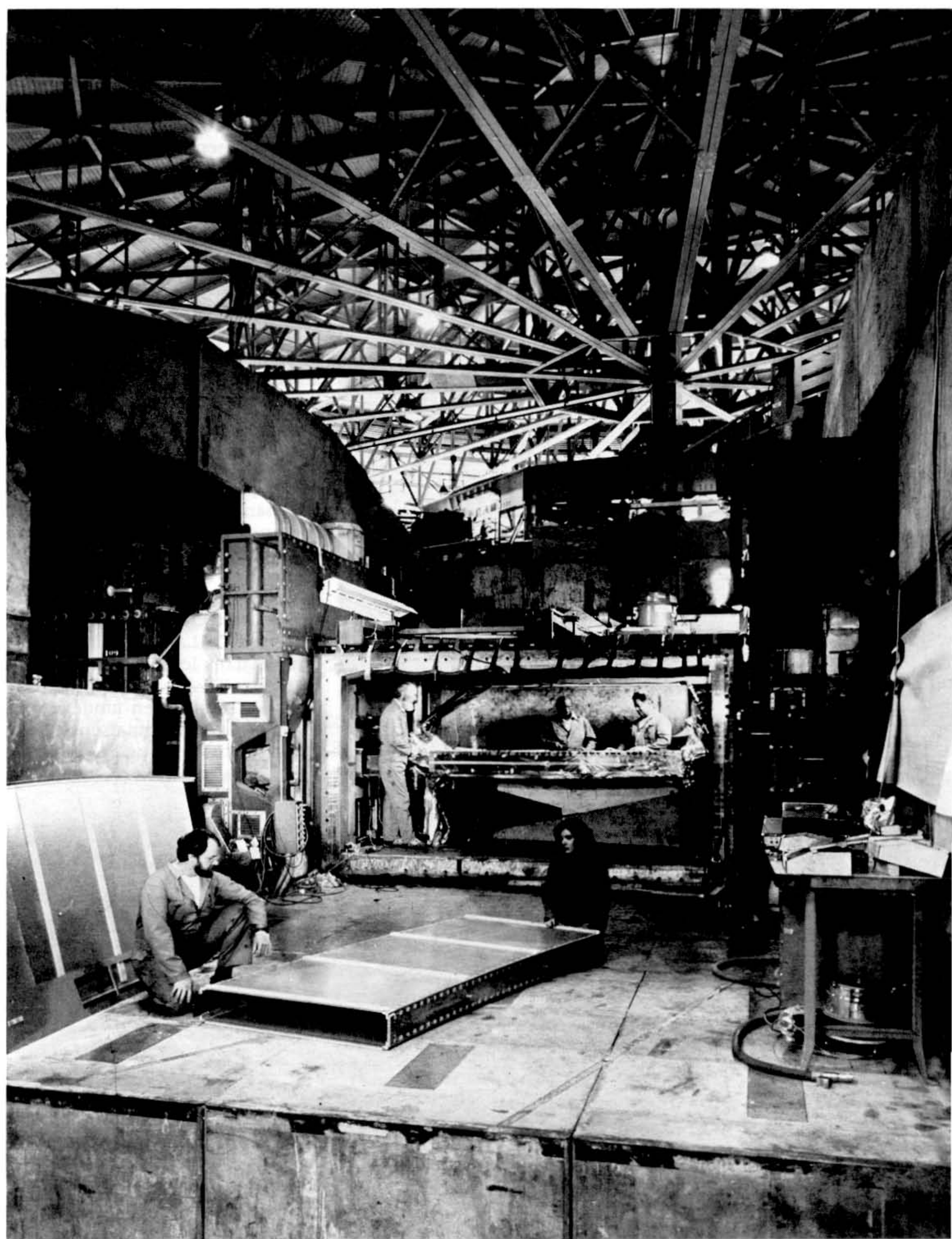


Planche III.20. Installation d'un nouveau tube à vide qui, par sections de trois mètres, a été placé dans l'intérieur de l'accélérateur un peu comme on insère des piles dans une lampe torche. Cette nouvelle enceinte à ultra-vide permet d'accélérer des ions d'uranium à des vitesses proches de celle de la lumière. (Avec l'aimable autorisation du laboratoire Lawrence Berkeley, université de Californie, Berkeley.)

– aussi longtemps que la Terre. Le seul danger qui en résulterait serait la *pollution thermique* – due à l'énergie de fusion s'ajoutant à la chaleur totale qui arrive à la surface de la Terre. Cela ferait légèrement monter la température et aurait des effets semblables à ceux de l'effet de serre. Cela serait également vrai de l'énergie solaire obtenue par une façon autre que la radiation solaire parvenant naturellement à la Terre. Par exemple, des stations solaires fonctionnant dans l'espace augmenteraient l'arrivée de chaleur à la surface du globe. Dans l'un ou l'autre cas, l'humanité devrait limiter sa consommation d'énergie ou mettre au point des méthodes pour rejeter la chaleur de la Terre dans l'espace à un taux supérieur au taux existant.

Cependant, tout ceci ne présente un intérêt que si la fusion nucléaire contrôlée peut être obtenue d'abord en laboratoire puis adaptée pour devenir un processus commercial pratique. Après une génération de travail, nous n'avons pas encore atteint ce stade.

Des trois isotopes de l'hydrogène, l'hydrogène 1 est le plus commun mais aussi le plus difficile à porter à la fusion. C'est le combustible du Soleil, mais celui-ci le contient par milliers de milliards de kilomètres cubes, d'où un gigantesque champ de gravitation pour le maintenir condensé, avec des températures centrales de plusieurs millions de degrés. Ce n'est qu'un minuscule pourcentage de l'hydrogène du Soleil qui subit la fusion à chaque instant ; mais étant donné l'énorme masse présente, un minuscule pourcentage suffit pour fournir une immense énergie.

L'hydrogène 3 est le plus facile à porter à la fusion, mais il existe en quantités si minimes et doit être fabriqué au prix d'une dépense d'énergie si élevée qu'il est inutile d'envisager à l'heure actuelle son emploi comme combustible.

Cela nous laisse l'hydrogène 2 ou deutérium, qui est plus facile à manier que l'hydrogène 1 et beaucoup plus commun que l'hydrogène 3. Dans tout l'hydrogène du globe il n'y a qu'un atome sur 6 000 qui soit du deutérium, mais cela suffit. Il y a 35 000 milliards de tonnes de deutérium dans les océans, assez pour fournir à l'homme de l'énergie pour tous les temps futurs.

Il subsiste cependant des problèmes. Ceci peut paraître surprenant, puisque les bombes à fusion existent. Si nous pouvons amener de l'hydrogène à la fusion, pourquoi ne pas fabriquer un réacteur aussi bien qu'une bombe ? Oui, mais pour faire une bombe à fusion, nous avons besoin d'utiliser une bombe à fission comme allumeur, puis nous laissons aller les choses. Pour faire un réacteur à fusion, nous avons besoin d'un allumeur plus doux, puis nous devons maintenir la réaction à une vitesse constante, contrôlée – et non explosive.

Le premier problème est le moins difficile. Des courants électriques forts, des ondes sonores de haute énergie, des faisceaux lasers, etc., peuvent tous produire très brièvement des températures de l'ordre du million de degrés. On ne doute pas de pouvoir atteindre les températures nécessaires.

Maintenir la température en gardant l'hydrogène en fusion en place (c'est ce qu'on espère) est tout autre chose. Bien évidemment, aucun matériau d'enceinte ne peut garder un gaz à plus de 100 millions de degrés. Ou l'enceinte serait vaporisée ou le gaz refroidirait. Le premier pas vers une solution est de réduire la densité du gaz bien au-dessous de celle à pression normale, diminuant ainsi le contenu en énergie, bien que l'énergie de chaque particule reste élevée. Le second pas est un concept d'une grande

ingéniosité. Dans un gaz à très haute température, les atomes ont perdu tous leurs électrons ; ils forment un *plasma* (un terme introduit par Irving Langmuir au début des années trente) fait d'électrons et de noyaux nus. Comme le plasma est entièrement constitué de particules chargées, pourquoi ne pas utiliser un fort champ magnétique, prenant la place de l'enceinte, pour le contenir ? Le fait que les champs magnétiques peuvent confiner des particules chargées et resserrer un faisceau de celles-ci est connu depuis 1907 ; on l'a appelé l'*effet de pincement ou de striction.* On avait essayé l'idée de la *bouteille magnétique* et cela avait marché – mais uniquement pendant des temps extrêmement courts (figure 10.6). Les filets de plasma pincés dans la bouteille se tordaient immédiatement comme un serpent, se cassaient et disparaissaient.

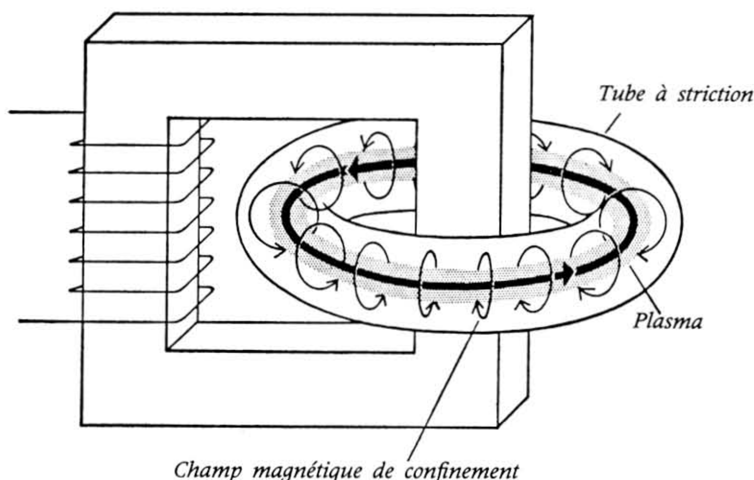


Figure 10.6. Bouteille magnétique conçue pour confiner un gaz chaud de noyaux d'hydrogène (un plasma). L'anneau est appelé *tore*.

Une autre approche consiste à appliquer un champ magnétique plus fort aux extrémités du tube, de façon à repousser le plasma et à l'empêcher ainsi de fuir. Ce n'est pas encore suffisant, bien que de fort peu. Si seulement un plasma à 100 millions de degrés pouvait tenir en place pendant environ une seconde, la réaction de fusion pourrait démarrer, et l'on extrairait de l'énergie du système. Cette énergie pourrait servir à renforcer le champ magnétique et à maintenir la température au niveau nécessaire. La réaction de fusion serait alors auto-entretenue, l'énergie produite servant à la maintenir. Mais empêcher le plasma de fuir ne serait-ce que pour cette petite seconde dépasse de beaucoup les possibilités actuelles. Comme la fuite du plasma se produit avec plus de facilité au bout du tube, pourquoi ne pas supprimer les extrémités en donnant au tube une forme en anneau ? Une conception particulièrement adaptée est celle du tube en anneau (tore), tordu en forme de huit. Le dispositif en forme de huit a été conçu pour la première fois en 1951 par Spitzer et

est appelé un *stellarator*. Un dispositif encore plus prometteur a été conçu par le physicien soviétique Lev Andreevitch Artsimovitch. On l'appelle *Toroidal Kamera Magnetic*, en abrégé *Tokamak*.

Les physiciens américains travaillent aussi sur des Tokamaks et, en outre, sur un dispositif appelé Scyllac, qui est conçu pour confiner un gaz plus dense, requérant donc une plus courte durée de confinement.

Pendant près de vingt ans, les physiciens ont avancé pas à pas vers l'utilisation de l'énergie de la fusion. Les progrès ont été lents, mais jusqu'ici rien ne laisse penser qu'ils conduiront à une impasse.

Entre-temps, il a fallu trouver des applications pratiques aux découvertes des chercheurs sur la fusion. Des torches à plasma émettant des jets à des températures allant jusqu'à 50 000 °C dans un silence absolu peuvent supplanter de loin les chalumeaux chimiques ordinaires. Et on a suggéré que la torche à plasma pourrait constituer l'unité d'élimination des déchets la plus parfaite. Dans sa flamme, n'importe quoi, véritablement *n'importe quoi* serait décomposé en ses éléments constitutifs, et tous ces éléments pourraient être à nouveau combinés entre eux et convertis en matériaux réutilisables.

SECONDE PARTIE

Les sciences biologiques

Chapitre 11

La molécule

La matière organique

Le mot *molécule* (diminutif du mot latin signifiant « masse ») indiquait à l'origine l'élément, l'unité indivisible d'une substance ; et dans un sens, *c'est* une particule élémentaire, car elle ne peut pas être cassée sans perdre son identité. Bien sûr, une molécule de sucre ou d'eau peut être divisée en atomes ou groupes d'atomes, mais ils ne sont plus alors ni sucre ni eau. Même une molécule d'hydrogène perd ses propriétés chimiques caractéristiques quand on la sépare en ses deux atomes d'hydrogène.

De même que l'atome a été le centre d'intérêt de la physique au XXe siècle, la molécule a été le sujet de découvertes tout aussi passionnantes en chimie. Les chimistes ont représenté de manière détaillée la structure de molécules très complexes, identifié le rôle de molécules spécifiques des systèmes vivants, créé de nouvelles molécules élaborées, prédit le comportement de la molécule d'une structure donnée avec une étonnante précision.

Dès le milieu du XXe siècle, les molécules complexes qui constituent les clefs de voûte du tissu vivant, les protéines et les acides nucléiques, étaient étudiées à l'aide de toutes les techniques rendues possibles par les progrès de la chimie et de la physique. Deux sciences : *la biochimie* (étude des réactions chimiques ayant lieu dans les tissus vivants) et la *biophysique* (étude des phénomènes et des forces physiques impliqués dans les processus vivants) ont fusionné pour donner naissance à une nouvelle discipline : *la biologie moléculaire*. Grâce aux découvertes de cette dernière,

la science moderne a presque effacé d'un seul trait la frontière entre le vivant et le non-vivant.

Il y a moins d'un demi-siècle, on ignorait encore la structure de la molécule la plus simple. Au début du XIXᵉ siècle, les chimistes n'étaient capables que de séparer la matière en deux grandes catégories. Ils savaient depuis longtemps (même du temps des alchimistes) que les substances sont classables en deux catégories nettement distinctes quant à leur réponse à la chaleur. Un groupe – par exemple le sel, le plomb, l'eau – reste pratiquement inchangé après avoir été chauffé. Le sel peut rougeoyer fortement quand il est chauffé, le plomb peut fondre, l'eau peut s'évaporer, mais quand ils sont ramenés à leur température de départ, ils reprennent leur forme première, apparemment sans aucun dommage. Le deuxième groupe de substances – par exemple le sucre, l'huile d'olive – est modifié de manière permanente par la chaleur. Le sucre se carbonise quand il est chauffé et reste carbonisé après avoir été refroidi ; l'huile d'olive se vaporise, et sa vapeur ne se condense pas après refroidissement. Les chercheurs ont remarqué que les substances résistant à la chaleur provenaient généralement du monde inanimé, de l'air, de l'Océan et du sol, tandis que les substances combustibles provenaient habituellement du monde de la vie, soit directement de la matière vivante, soit de ses restes après la mort. En 1807 Berzelius, qui inventa les symboles chimiques et devait préparer la première liste de poids moléculaires adéquate (voir chapitre 6) appela *organiques* les substances combustibles (parce qu'elles sont dérivées, directement ou indirectement, des organismes vivants) et tout le reste *inorganique*, ou *minéral*.

Au début, la chimie s'est surtout intéressée aux substances minérales. C'est l'étude du comportement des gaz inorganiques qui a conduit au développement de la théorie atomique. Cette théorie, une fois établie, a permis de clarifier rapidement la nature des molécules inorganiques. L'analyse montra que ces molécules sont en général constituées d'un petit nombre d'atomes différents, présents en proportions variables. La molécule d'eau contient deux atomes d'hydrogène et un atome d'oxygène ; la molécule de sel contient un atome de sodium et un atome de chlore ; l'acide sulfurique contient deux atomes d'hydrogène, un de soufre et quatre d'oxygène, et ainsi de suite.

Quand les chimistes ont commencé à analyser les substances organiques, le tableau se présentait autrement. Deux substances pouvaient avoir exactement la même composition tout en présentant des propriétés nettement différentes. Par exemple, l'alcool éthylique est composé de deux atomes de carbone, un d'oxygène et six d'hydrogène ; il en est de même de l'éther diméthylique – pourtant l'un est un liquide à la température ambiante, alors que l'autre est un gaz. Les molécules organiques contiennent plus d'atomes que les simples molécules inorganiques. Il n'y a apparemment pas de logique dans l'arrangement de leurs atomes. Les composés organiques ne pouvaient pas être expliqués à l'aide des lois élémentaires de la chimie qui s'appliquaient si facilement aux substances minérales.

Berzelius déclara que la chimie de la vie était quelque chose à part, qui obéissait à son propre jeu de règles subtiles. Seul le tissu vivant, dit-il, pouvait donner naissance à un composé organique. Son point de vue est un exemple de *vitalisme*.

C'est alors, en 1828, que le chimiste allemand Friedrich Wöhler, un étudiant de Berzelius, fabriqua une substance organique dans son

laboratoire ! Il était en train de chauffer un composé, le *cyanate d'ammonium*, généralement considéré comme inorganique, lorsqu'il constata, à sa stupéfaction, que ce matériau se transformait en un composé blanc aux propriétés identiques à celles de *l'urée*, un composant de l'urine. Selon le point de vue de Berzelius, seul du tissu vivant pouvait produire de l'urée ; or, Wöhler en avait produit à partir d'un composé minéral simplement en le chauffant.

Wöhler répéta de nombreuses fois l'expérience avant de se risquer à la publier. Quand finalement il s'y décida, Berzelius et bien d'autres refusèrent de le croire, jusqu'à ce que des chimistes confirment ses résultats. De plus, ces derniers réussirent à synthétiser de nombreux composés organiques à partir de produits inorganiques. Le premier à produire un composé organique à partir d'éléments simples fut le chimiste allemand Adolph Wilhelm Hermann Kolbe, qui fabriqua ainsi, en 1845, de *l'acide acétique* (la substance qui donne son goût au vinaigre). Ce fut ce travail qui détruisit vraiment le « vitalisme » de Berzelius. Il devint évident que les mêmes lois chimiques s'appliquaient aussi bien aux molécules organiques qu'aux molécules inorganiques. Finalement, on arriva à une définition simple de la distinction entre substances organiques et minérales : toutes les substances contenant du carbone (à l'exception possible de quelques composés simples, tels que le gaz carbonique) sont appelées organiques ; les autres sont minérales.

LA STRUCTURE CHIMIQUE

Les chimistes avaient besoin d'un système de notation simple pour représenter les composés et utiliser cette nouvelle chimie. Berzelius avait déjà suggéré un système de symboles rationnel et pratique dans lequel les éléments sont désignés par l'abréviation de leur nom latin : *C* représente le carbone, *O* l'oxygène, *H* l'hydrogène, *N* l'azote, *S* le soufre, *P* le phosphore, etc. Quand deux éléments commencent par la même lettre, une deuxième lettre est utilisée pour les distinguer : par exemple, *Ca* pour le calcium, *Cl* pour le chlorure, *Cd* pour le cadmium, *Co* pour le cobalt, *Cr* pour le chrome, et ainsi de suite. Dans quelques rares cas, les noms latins ou latinisés (et leurs initiales) diffèrent du français, tels que or (*aurum*) : *Au*, étain (*stannum*) : *Sn*, mercure (*hydragyrum*) : *Hg*, antimoine (*stibium*) : *Sb*, sodium (*natrium*) : *Na*, et potassium (*kalium*) : *K*.

Il est facile de représenter la composition d'une molécule à l'aide de ce système. L'eau est écrite H_2O (indiquant donc que la molécule est constituée de deux atomes d'hydrogène et d'un atome d'oxygène) ; le sel, NaCl ; l'acide sulfurique, H_2S0_4, etc. Ceci constitue la *formule brute* d'un composé ; elle indique de quoi il est fait, mais ne dit rien sur sa structure – c'est-à-dire sur la manière dont ses atomes sont liés entre eux.

En 1831, le baron Justus von Liebig, un collaborateur de Wöhler, commença à étudier la composition d'un certain nombre de composés organiques, appliquant l'*analyse chimique* à la chimie organique. Il décomposait par combustion une petite quantité de substance organique ; les gaz ainsi formés (principalement du gaz carbonique, CO_2, et de la vapeur d'eau, H_2O), étaient captés à l'aide des produits chimiques appropriés, qu'il pesait ensuite pour en mesurer l'augmentation en poids due aux gaz piégés. De ce poids il pouvait déduire la quantité de carbone,

d'hydrogène et d'oxygène présente dans la substance originelle. Il était alors facile de calculer, à partir du poids moléculaire, le nombre de chaque type d'atome dans la molécule. Liebig établit ainsi que la formule de la molécule d'alcool éthylique est C_2H_6O, par exemple.

La méthode de Liebig ne permettait pas de mesurer la quantité d'azote présente dans un composé organique ; en 1833, le chimiste français Jean-Baptiste André Dumas mit au point une méthode de combustion grâce à laquelle on pouvait collecter l'azote gazeux libéré par la substance. Il utilisa cette méthode, en 1841, pour analyser la composition gazeuse de l'atmosphère avec une précision sans précédent.

Les techniques d'*analyse de chimie organique* sont devenues de plus en plus sophistiquées jusqu'à atteindre les prodiges de raffinement des méthodes de *micro-analyse* du chimiste autrichien Fritz Pregl, mises au point en 1909, pour analyser avec précision des substances organiques en quantités à peine visibles à l'œil nu, ce qui lui valut le prix Nobel de chimie en 1923.

Malheureusement, on ne peut pas comprendre la chimie des composés organiques à l'aide de leur seule formule brute. Au contraire des composés minéraux, qui ne sont en général constitués que de deux ou trois atomes, ou tout au plus d'une douzaine, les molécules organiques sont souvent énormes. Liebig trouva que la formule de la morphine était $C_{17}H_{19}0_3N$, et celle de la strychnine, $C_{21}H_{22}O_2N_2$.

Les chimistes étaient perplexes devant des molécules aussi grandes dont la formule n'avait ni queue ni tête. Wöhler et Liebig essayèrent de regrouper les atomes en des sous-ensembles plus petits qu'ils appelèrent *radicaux.* Ils élaborèrent des théories pour démontrer que ces composés variés étaient faits de radicaux différents en nombre et en composition. Il était particulièrement difficile d'expliquer pourquoi deux composés, tels que l'alcool éthylique et l'éther diméthylique, ayant la même formule brute, avaient des propriétés différentes.

Ce problème fut mis en évidence autour des années 1820 par Liebig et Wöhler. Le premier étudiait un groupe de composés appelés les *fulminates ;* le second, un groupe appelé les *isocyanates* – et les deux s'avérèrent avoir la même formule brute. Les éléments y étaient présents dans les mêmes proportions. Ils parlèrent de cette découverte à Berzelius, le dictateur chimiste du moment, qui fut peu disposé à les croire, jusqu'à ce qu'il parvienne lui aussi à la même découverte, vers 1830. Il donna le nom d'*isomères* (du mot grec signifiant « parties égales ») aux composés ayant des propriétés différentes, mais possédant les mêmes éléments en proportions identiques. La structure des molécules organiques était un véritable puzzle à cette époque.

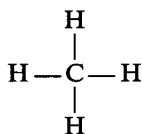
Les chimistes, perdus dans la jungle de la chimie organique, commencèrent à voir la lumière au bout du tunnel dans les années 1850, quand ils remarquèrent que chaque atome ne pouvait se combiner qu'avec un nombre donné d'autres atomes. Par exemple, ils constatèrent que l'atome d'hydrogène ne peut se lier qu'à un seul atome : il peut former l'acide chlorhydrique, HCl, mais jamais HCl_2. De même, le chlore et le sodium peuvent chacun prendre un seul partenaire, formant ainsi NaCl. Par contre, un atome d'oxygène peut s'apparier au maximum à deux atomes, par exemple H_2O. L'azote peut s'apparier au maximum à trois atomes, par exemple NH_3 (l'ammoniac). Et à l'atome de carbone peuvent se lier jusqu'à quatre atomes, par exemple CCl_4 (le tétrachlorure de carbone).

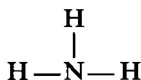
En bref, il semblait que chaque type d'atome possédait un certain nombre de crochets auxquels puissent se fixer d'autres atomes. Le chimiste anglais Edward Frankland, en 1852, fut le premier à exprimer ceci clairement, en appelant ces crochets liaisons de *valence,* du mot latin signifiant « valeur », c'est-à-dire le nombre indiquant la capacité de combinaison des éléments.

Le chimiste allemand Friedrich August Kekule von Stradonitz remarqua que si l'on donnait une valence de 4 au carbone, et si l'on émettait l'hypothèse que les atomes de carbone peuvent utiliser au moins une partie de ces valences pour se lier en chaînes, alors on pourrait se retrouver dans la jungle de la chimie organique. Son hypothèse fut rendue plus claire par la suggestion du chimiste écossais, Archibald Scott Couper, de représenter par des tirets ces forces de combinaison entre atomes (*liaisons,* selon leur appellation habituelle). De cette façon, les molécules organiques pouvaient être construites comme un jeu de « Meccano ».

En 1861, Kekule publia un livre contenant de nombreux exemples de son système, qui s'avéra pratique et utile. La *formule développée* devint l'estampille du chimiste organicien.

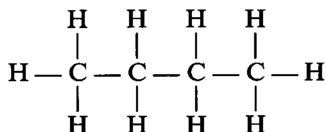
Par exemple, les molécules de méthane (CH_4), d'ammoniac (NH_3), et d'eau (H_2O) peuvent être respectivement représentées de la façon suivante :



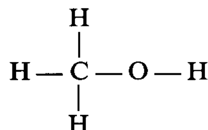


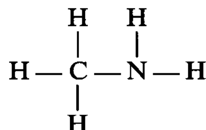


On peut représenter les molécules organiques par des chaînes d'atomes de carbone sur les côtés desquelles sont attachés des atomes d'hydrogène. En conséquence, la structure du butane (C_4H_{10}) est :

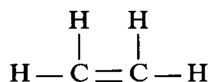


Les structures des deux composés : l'alcool méthylique (CH_4O) et la méthylamine (CH_5N), servent respectivement à montrer comment l'oxygène et l'azote peuvent entrer dans la chaîne.



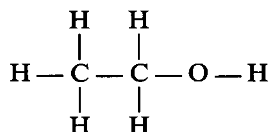


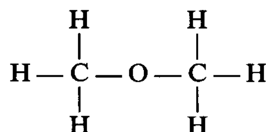
Un atome possédant plus d'un crochet, comme le carbone qui en a quatre, n'est pas obligé d'utiliser chacun d'entre eux pour se lier avec des atomes différents : il peut former une liaison double ou triple avec ses voisins, tel que dans l'éthylène (C_2H_4) ou l'acétylène (C_2H_2) :





Il devient maintenant facile de voir comment deux molécules peuvent avoir le même nombre d'atomes de chaque élément et cependant présenter des propriétés différentes. Les deux isomères doivent différer dans l'arrangement de leurs atomes. Par exemple, les formules développées de l'alcool éthylique et de l'éther diméthylique s'écrivent respectivement :





Plus le nombre d'atomes dans la molécule est grand, plus le nombre d'arrangements possibles est grand et plus le nombre d'isomères est grand. Par exemple, la molécule d'heptane, constituée de sept atomes de carbone et de seize atomes d'hydrogène, peut être arrangée de neuf façons différentes ; en d'autres termes, il peut y avoir neuf heptanes différents, chacun avec des propriétés particulières. Ces neuf isomères se ressemblent beaucoup, mais c'est seulement une ressemblance familiale. Les chimistes ont préparé ces neuf isomères mais n'ont jamais pu en trouver un dixième – une bonne preuve en faveur du système de Kekule.

Un composé contenant quarante atomes de carbone et quatre-vingt-deux atomes d'hydrogène peut être combiné de 62,5 millions de manières différentes, ou isomères. Des molécules organiques de cette taille ne sont nullement rares.

Pratiquement, seul le carbone peut former de longues chaînes. Les autres atomes arrivent au mieux à former une chaîne d'une demi-douzaine d'atomes environ. C'est pourquoi les molécules inorganiques sont habituellement simples et n'ont que rarement des isomères. La grande complexité des molécules organiques, liée aux possibilités d'isomérisation, fait que des millions de composés organiques sont connus, que de nouveaux sont créés chaque jour, et qu'un nombre à peu près infini d'entre eux attend d'être découvert.

Les formules développées sont universellement utilisées depuis 1861, car elles permettent de mettre en évidence la nature des molécules organiques. Les chimistes écrivent souvent la formule d'une molécule, pour la simplifier, en regroupant les atomes la constituant sous forme de radicaux, tels que les radicaux méthyle (CH_3) et méthylène (CH_2). La formule du butane peut donc être écrite : $CH_3CH_2CH_2CH_3$.

Les détails de la structure

Pendant la deuxième moitié du XIXe siècle, les chimistes ont mis en évidence une forme d'isomérisme particulièrement subtile qui s'est avérée très importante dans la chimie du monde vivant. La découverte provient d'un effet asymétrique que certains composés organiques ont sur les rayons lumineux qui les traversent.

L'ACTIVITÉ OPTIQUE

Dans un plan perpendiculaire à la direction d'un faisceau de lumière ordinaire, les ondes innombrables qui le composent vibrent dans toutes les directions du plan – vers le haut, vers le bas, d'un côté à l'autre, et obliquement. Une telle lumière est *non polarisée*. Mais quand la lumière passe, par exemple, dans le cristal d'une substance transparente appelée *spath d'Islande*, elle est réfractée de façon telle qu'elle émerge *polarisée* dans un plan. Tout se passe comme si seuls certains plans de vibration pouvaient traverser le réseau d'atomes du cristal ; tout comme les pieux d'une clôture permettent à une personne se déplaçant de côté de se frayer un passage, mais pas à une personne arrivant de face. Il existe des instruments tels que le *prisme de Nicol*, mis au point par le physicien écossais William Nicol en 1829, qui laissent passer la lumière dans un seul plan (figure 11.1).

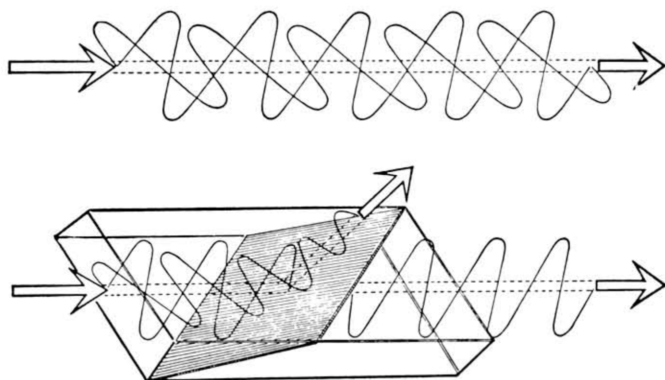


Figure 11.1. La polarisation de la lumière. Les ondes lumineuses oscillent normalement dans tous les plans (*en haut*). Le prisme de Nicol (*en bas*) ne laisse passer les oscillations que dans un seul plan, réfléchissant les autres rayons dans d'autres directions. La lumière transmise est polarisée dans un plan.

Le prisme a maintenant été remplacé dans les polarimètres, dans la plupart des cas, par des polaroïds (cristaux d'un complexe de sulfate de quinine et d'iode, alignés suivant des axes parallèles et noyés dans la nitrocellulose), mis au point en premier par Edwin Land en 1932.

La lumière réfléchie est souvent partiellement polarisée dans un plan, découverte faite en 1808 par le physicien français Étienne Louis Malus. Il inventa le terme *polarisation* à partir d'une remarque de Newton, d'ailleurs inexacte, sur les pôles des particules lumineuses. L'éclat de la lumière réfléchie sur les fenêtres des immeubles, des voitures ainsi que sur les routes pavées peut être ramené à des niveaux supportables par l'utilisation de lunettes polaroïds.

En 1815, le physicien français Jean-Baptiste Biot découvrit que lorsqu'un faisceau de lumière polarisée dans un plan passe à travers des cristaux de quartz, le plan de polarisation subit une rotation ; c'est-à-dire que la

lumière entre en vibrant dans un plan et sort en vibrant dans un plan différent. On dit d'une substance qui agit ainsi qu'elle a une *activité optique.* Certains cristaux dévient le plan de la lumière polarisée dans le sens des aiguilles d'une montre (*dextrorotation*) et d'autres dans le sens contraire (*lévorotation*). Biot observa que des substances organiques, telles que le camphre ou l'acide tartrique, agissent ainsi. Il pensait qu'il était vraisemblable qu'une certaine asymétrie dans l'arrangement des atomes soit responsable de la rotation de la lumière. Mais cette suggestion resta une pure spéculation pendant longtemps.

En 1844, Louis Pasteur, à l'âge de vingt-deux ans, s'attaqua à cette question intéressante. Il étudia deux substances, l'acide tartrique et l'acide racémique, de même composition, avec la différence que l'acide tartrique dévie le plan de la lumière polarisée, tandis que l'acide racémique ne le dévie pas. Pasteur pensait que les cristaux des sels de l'acide tartrique s'avéreraient asymétriques et que ceux de l'acide racémique seraient symétriques. A sa grande surprise, en examinant les deux types de cristaux au microscope, il découvrit qu'ils étaient tous les deux asymétriques. En fait, il y avait deux types de cristaux de racémate : la moitié d'entre eux avaient la même forme que les cristaux du tartrate, et l'autre moitié était leur image dans un miroir. En quelque sorte, la moitié des cristaux de racémate était gauche et l'autre moitié était droite.

Pasteur sépara patiemment les cristaux de racémate gauches des droits, puis il fit dissoudre chacun d'eux séparément et passer la lumière à travers chaque solution. La solution de cristal ayant la même asymétrie que les cristaux de tartrate déviait la lumière polarisée exactement comme le fait le tartrate, dans le même sens et avec la même amplitude. Ces cristaux *étaient* du tartrate. L'autre solution déviait le plan de la lumière polarisée avec la même amplitude de rotation, mais dans le sens opposé. Le racémate de départ ne déviait pas la lumière polarisée car les tendances opposées des cristaux le constituant s'annulaient l'une l'autre.

Ensuite, Pasteur reconvertit en acide les deux *sels* de racémate qu'il avait séparés, par addition d'ions hydrogène (un *sel* est un composé dans lequel certains hydrogènes de la molécule d'acide sont remplacés par d'autres ions chargés positivement, tels que le sodium ou le potassium). Il trouva que ces deux acides racémiques étaient optiquement actifs – l'un déviant la lumière polarisée dans le même sens que le faisait l'acide tartrique (parce que c'était de l'acide tartrique), et l'autre dans la direction opposée.

On découvrit d'autres couples de composés symétriques par rapport à un miroir (*éniantiomorphes,* du grec signifiant « formes opposées »). En 1863, le chimiste allemand Johannes Wislicenus trouva que l'acide lactique (l'acide du petit lait) formait ce genre de couple. De plus, il montra que les propriétés des deux éléments du couple étaient identiques, *à part* la déviation de la lumière polarisée. On vérifia cette propriété pour tous les éniantiomorphes.

Jusqu'ici tout va bien, mais d'où provenait l'asymétrie ? Qu'y avait-il dans ces molécules qui en faisait des images symétriques par rapport à un miroir ? Pasteur ne pouvait le dire. Et bien que Biot en ait eu l'intuition, il mourut à quatre-vingt-huit ans, avant de voir la démonstration de l'existence d'une asymétrie moléculaire.

La réponse fut finalement trouvée en 1874, douze ans après la mort de Biot. Deux jeunes chercheurs, un Néerlandais de vingt-deux ans du nom de Jacobus Hendricus Van't Hoff, et un Français de vingt-sept ans

nommé Joseph Achille Le Bel, proposèrent chacun de leur côté une nouvelle théorie des liaisons de valence du carbone, qui expliquait comment on pouvait obtenir des molécules symétriques par rapport à un miroir. Plus tard, Van't Hoff étudia les substances en solution et montra que les mêmes lois régissent le comportement des solutions et des gaz. Pour cette découverte, il fut le premier, en 1901, à recevoir le prix Nobel de chimie.

Kekule avait dessiné les quatre liaisons de l'atome de carbone dans le même plan, pas nécessairement parce que c'était ainsi qu'elles étaient réellement arrangées, mais parce que c'était une façon commode de les dessiner sur une feuille de papier. Van't Hoff et Le Bel proposèrent un modèle tridimensionnel dans lequel les liaisons étaient dirigées dans deux plans perpendiculaires, deux dans un plan et deux dans l'autre. Une bonne manière de représenter ce modèle est d'imaginer l'atome de carbone reposant sur trois pieds, le quatrième pointant vers le haut. Si on suppose que l'atome de carbone est au centre d'un *tétrahèdre* (une pyramide régulière à quatre côtés, ayant des côtés triangulaires), alors les quatre liaisons pointent vers les quatre sommets de la figure. Le modèle est donc appelé *modèle de l'atome de carbone tétrahédrique* (voir figure 11.2).

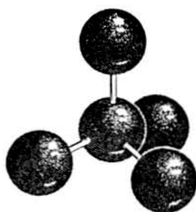


Figure 11.2. Le carbone tétrahédrique.

Maintenant, attachons à ces quatre liaisons deux atomes d'hydrogène, un atome de chlore, et un atome de brome. Quel que soit l'ordre de liaison, nous arrivons toujours au même arrangement. Essayez pour voir. Vous pouvez représenter les quatre liaisons en piquant quatre cure-dents dans une boule de guimauve (l'atome de carbone) avec des angles corrects. Maintenant, supposez que vous piquez, dans n'importe quel ordre, deux olives noires (les atomes d'hydrogène), une olive verte (le chlore), et une cerise (le brome) au bout de chaque cure-dent. Disons que quand vous faites tenir cet assemblage sur trois pieds avec une olive noire pointant vers le haut, l'ordre des trois pieds, dans le sens des aiguilles d'une montre, est olive noire, olive verte, et cerise. Vous pouvez maintenant intervertir l'olive verte et la cerise de telle sorte que l'ordre soit olive noire, cerise, olive verte. Pour retrouver le même ordre que précédemment, tout ce que vous avez à faire est de retourner la structure de façon que l'olive noire, qui servait de pied, pointe en l'air, et que celle qui était en l'air soit posée sur la table. L'ordre des pieds est alors de nouveau olive noire, olive verte, cerise.

En d'autres termes, quand au moins deux des atomes (ou groupes d'atomes) liés aux quatre liaisons du carbone sont identiques, un seul arrangement structural est possible. Évidemment, c'est aussi vrai quand trois ou quatre des atomes liés sont identiques.

En revanche, quand les quatre atomes (ou groupes d'atomes) sont différents, la situation change. Deux arrangements structuraux sont alors possibles – l'un étant symétrique de l'autre par rapport à un miroir. Par exemple, supposons que vous plantez une cerise sur la pointe en l'air du carbone et une olive noire, une olive verte et un petit oignon sur les trois pointes qui forment la base du tétrahèdre. Si vous échangez l'olive noire et l'olive verte de telle sorte que l'ordre, dans le sens des aiguilles d'une montre, soit olive verte, olive noire, oignon, il n'y a aucun moyen de tourner la structure pour retrouver l'ordre olive noire, olive verte, oignon, tel qu'il était avant que vous ne fassiez l'échange. Donc, avec quatre liaisons différentes, vous pouvez toujours former deux structures différentes, images dans un miroir l'une de l'autre. Essayez et constatez.

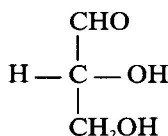
Van't Hoff et Le Bel éclaircirent ainsi le mystère de l'asymétrie des substances optiquement actives. Les substances symétriques par rapport à un miroir qui déviaient la lumière polarisée dans des directions opposées étaient des substances contenant au moins un atome de carbone ayant quatre atomes ou groupes d'atomes différents attachés à ses liaisons. Un des deux arrangements possibles de ces liaisons dévie la lumière vers la droite ; l'autre la dévie vers la gauche.

Le modèle du carbone tétrahédrique de Van't Hoff et Le Bel, étayé par des preuves de plus en plus nombreuses, et avec l'appui enthousiaste du célèbre Wislicenus, fut universellement accepté en 1885.

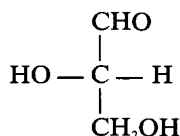
La tridimensionnalité fut appliquée à la représentation d'autres atomes que le carbone. Le chimiste allemand Viktor Meyer l'appliqua avec succès à l'atome d'azote, tandis que le chimiste anglais William Jackson Pope l'élargissait au soufre, au sélénium et à l'étain. Le chimiste suisse allemand Alfred Werner ajouta d'autres éléments à cette liste et, au début des années 1890, il élabora une *théorie de la coordination* expliquant la structure des substances minérales complexes par une étude approfondie de la distribution des atomes et groupes d'atomes autour d'un atome central ; il reçut le prix Nobel de chimie en 1913.

Les deux acides racémiques que Pasteur avait isolés furent appelés « d-acide tartrique » (pour « dextrogyre ») et « l-acide tartrique » (pour « lévogyre »). On écrivit les formules développées de ces composés symétriques par rapport à un miroir. Mais comment s'y retrouver ? Quel composé était réellement le droit et lequel le gauche ? Il était alors impossible de le dire.

Il fallait aux chimistes une référence, ou un étalon de comparaison, pour distinguer les substances gauches des substances droites. Le chimiste allemand Emil Fischer choisit un composé simple appelé le *glycéraldéhyde*, de la famille des sucres, qui était un des composés optiquement actifs les plus étudiés. Il décida arbitrairement qu'une des formes était gauche – il l'appela « L-glycéraldéhyde » –, et que son image dans un miroir était droite – il l'appela « D-glycéraldéhyde ». Leurs formules développées sont :



D-glycéraldéhyde



L-glycéraldéhyde

Tout composé dont on a pu démontrer par des méthodes chimiques appropriées qu'il a une structure apparentée au L-glycéraldéhyde, est considéré de la série « L » et a le préfixe « L » lié à son nom, qu'il soit lévogyre ou dextrogyre en ce qui concerne son effet sur la lumière polarisée. En l'occurrence, la forme lévogyre de l'acide tartrique appartient à la série D au lieu de la série L. De nos jours, le nom d'un composé qui appartient à la série D par sa structure, mais dévie la lumière polarisée vers la gauche, est précédé de « D(–) » ; de la même manière, il y a « D(+) », « L(–) », et « L(+) ».

Cette préoccupation concernant les causes de l'activité optique s'avéra être plus qu'une curiosité futile. En fait, presque tous les composés trouvés dans les organismes vivants contiennent des carbones asymétriques. Et dans chaque cas, l'organisme n'utilise que l'une des deux formes asymétriques du composé. De plus, les composés similaires sont généralement de la même série. Par exemple, presque tous les sucres simples trouvés dans les tissus vivants appartiennent à la série D, tandis que virtuellement tous les acides aminés (les chaînons des protéines) appartiennent à la série L.

En 1955, un chimiste hollandais nommé Johannes Martin Bijvoet établit définitivement quelle structure dévie la lumière polarisée vers la gauche, et inversement. Il s'avéra que Fischer avait, par hasard, désigné correctement les formes lévogyres et dextrogyres.

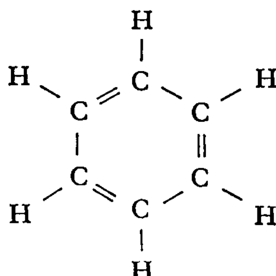
LE PARADOXE DE L'ANNEAU BENZÉNIQUE

Bien que Kekule ait établi de façon sûre son système de formule développée, il ne put, pendant plusieurs années, formuler une molécule relativement simple, le benzène, découverte en 1825 par Faraday. L'analyse élémentaire montrait qu'il était composé de six atomes de carbone et de six atomes d'hydrogène. Qu'était-il arrivé à toutes les liaisons de carbone en trop ? (Six atomes de carbone liés entre eux par des liaisons simples peuvent encore se lier à quatorze atomes d'hydrogène, comme dans l'hexane, C_6H_{14}). Évidemment, dans le benzène, les atomes de carbone étaient liés entre eux par des liaisons doubles ou triples. Le benzène pouvait donc avoir une formule telle que $CH \equiv C - CH = CH - CH = CH_2$. Mais le problème était que tous les composés connus ayant ce genre de structure avaient des propriétés très différentes de celles du benzène. En outre, toutes les évidences chimiques semblaient indiquer que la molécule de benzène était très symétrique, et six carbones et six hydrogènes ne pouvaient en aucune manière être arrangés en une chaîne raisonnablement symétrique.

En 1865, Kekule trouva la réponse. Il raconta quelques années plus tard que la vision de la molécule de benzène lui est apparue tandis qu'il était dans un autobus, plongé dans une rêverie, à demi endormi. Les chaînes d'atomes de carbone semblaient vivre et danser devant ses yeux, et soudainement, l'une d'entre elles s'enroula sur elle-même comme un

serpent. Kekule sortit de sa rêverie en sursaut et faillit crier « Eurêka » ! Il tenait la solution : le benzène était un cycle.

Kekule supposa que les six atomes de carbone de la molécule étaient arrangés de la façon suivante :



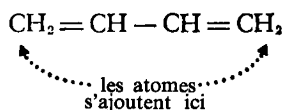
Là, nous avions enfin la symétrie requise. La structure expliquait, entre autres choses, pourquoi la substitution d'un atome d'hydrogène du benzène par un autre atome donnait toujours un seul et même produit ; tous les carbones de l'anneau étant équivalents en termes structuraux. Quel que soit l'endroit où la substitution est faite sur le cycle benzénique, on obtient le même produit. Deuxièmement, la structure cyclique montre qu'il y a seulement trois façons de remplacer deux atomes d'hydrogène du cycle : on peut faire la substitution soit sur deux carbones adjacents du cycle, soit sur deux atomes séparés par une seule liaison, soit sur deux carbones séparés par deux liaisons. Ainsi, on ne trouve que trois isomères du benzène doublement substitués.

Le schéma de la molécule de benzène proposé par Kekule posait cependant une question embarrassante. En général, les composés ayant des liaisons doubles sont plus réactifs, c'est-à-dire moins stables, que ceux qui n'ont que des liaisons simples. Comme si la liaison supplémentaire avait plus tendance à déserter l'attachement de l'atome de carbone pour former une nouvelle liaison. On peut très facilement fixer de l'hydrogène ou d'autres atomes aux composés à double liaison ; on peut aussi les scinder sans difficulté. Or, le cycle benzénique est extraordinairement stable, plus stable même que des chaînes de carbone ne comportant que des liaisons simples. En fait, il est si stable et si courant dans la matière organique que les molécules contenant un cycle benzénique constituent une grande classe de composés organiques : les composés *aromatiques*, tous les autres étant regroupés sous le nom d'*aliphatiques*. La molécule de benzène résiste à la fixation de molécules d'hydrogène supplémentaires et elle est difficile à rompre.

Les chimistes du XIXe siècle furent incapables de trouver une explication à cette étrange stabilité des doubles liaisons de la molécule de benzène ; tout le système de formules développées de Kekule était remis en question par la molécule de benzène. Leur incapacité à résoudre ce problème rendait tout le reste incertain.

Le chimiste allemand Johannes Thiele fut celui qui s'approcha le plus de la solution. En 1899, il émit l'hypothèse que, lorsque des liaisons simples et doubles alternent, les extrémités les plus rapprochées d'une paire de doubles liaisons se neutralisent et annulent leur réactivité réciproque. Considérons par exemple un composé, le *butadiène*, le modèle le plus simple de deux doubles liaisons séparées par une simple liaison (*des doubles liaisons*

conjuguées). Si on ajoute deux atomes à ce composé, ils se lient aux carbones terminaux, comme le montre la formule ci-dessous. Cette hypothèse expliquait la non-réactivité du benzène. Ses trois doubles liaisons étant arrangées dans un cycle, elles se neutralisent complètement les unes les autres.



Quelque quarante ans plus tard, la nouvelle théorie de la liaison chimique, qui représentait les liaisons par la mise en commun d'électrons (voir chapitre 6), apporta une réponse plus satisfaisante. Chaque atome qui se combine avec un partenaire partage un de ses électrons avec ce dernier, et le partenaire donne un de ses électrons à la liaison. Le carbone, avec quatre électrons dans sa couche extérieure, peut former quatre liaisons ; l'hydrogène met en commun son seul électron pour une liaison avec un autre atome, etc.

La question soulevée est la suivante : comment les électrons sont-ils répartis ? Évidemment, deux atomes de carbone se partagent également entre eux la paire d'électrons, puisque chaque atome retient les électrons de façon équivalente. En revanche, dans une combinaison telle que H_2O, l'atome d'oxygène, qui retient plus fortement les électrons que l'atome d'hydrogène, accapare une plus grande part des électrons qu'il a en commun avec chaque atome d'hydrogène. Donc, l'atome d'oxygène, en vertu de sa part plus grande d'électrons, a un léger excès de charge négative. Du même coup, l'atome d'hydrogène, souffrant d'un manque d'électrons, a une charge légèrement positive. Une molécule ayant un appariement oxygène-hydrogène, comme l'eau ou l'alcool éthylique, possède une faible concentration de charge négative dans une partie de la molécule et une faible concentration de charge positive dans une autre. Cette molécule possède donc deux pôles de charge, et elle est appelée *polaire*.

Cette conception de la structure de la molécule fut proposée en premier en 1912 par Peter Debye (qui suggéra plus tard une méthode basée sur le magnétisme pour atteindre des températures très basses; voir chapitre 6). Il utilisa un champ électrique pour mesurer la quantité de charge séparant les pôles électriques d'une molécule. Dans un tel champ, les molécules polaires s'alignent, orientant leurs extrémités négatives vers le pôle positif du champ et leurs extrémités positives vers le pôle négatif. La facilité avec laquelle cet alignement se fait est une mesure du *moment dipolaire* de la molécule. Dès le début des années trente, la mesure du moment dipolaire des molécules était devenue pratique courante ; et en 1936, pour ce travail, entre autres, Debye reçut le prix Nobel de chimie.

Cette théorie de la polarité rendait compte de certains phénomènes jusqu'alors inexpliqués, les anomalies du point d'ébullition des substances, par exemple. En général, plus le poids moléculaire d'un composé est grand, plus son point d'ébullition est élevé. Mais il y a de nombreuses exceptions à cette règle. L'eau, dont le poids moléculaire n'est que de 18, bout à 100 °C, tandis que le propane, ayant plus de deux fois ce poids moléculaire (44), bout à une température beaucoup plus basse de – 42 °C. Pourquoi cette

différence ? La réponse est que l'eau est une molécule polaire ayant un moment dipolaire élevé, tandis que le propane est *apolaire* – il n'a pas de pôles de charge. Les molécules polaires ont tendance à s'orienter de telle manière que le pôle négatif d'une molécule soit adjacent au pôle positif de sa voisine. L'attraction qui en résulte entre molécules voisines rend leur séparation plus difficile, et ces substances ont un point d'ébullition relativement élevé. En conséquence, l'alcool éthylique a un point d'ébullition beaucoup plus haut (78 °C) que son isomère l'éther diméthylique, qui bout à – 24 °C, bien que ces substances aient le même poids moléculaire (46). L'alcool éthylique a un grand moment dipolaire, alors que celui de l'éther diméthylique est faible. L'eau a un moment dipolaire encore plus grand que celui de l'alcool éthylique.

Quand de Broglie et Schrödinger formulèrent une nouvelle théorie relative aux électrons, considérés non plus comme des particules précisément définies mais comme des paquets d'ondes (voir chapitre 8), la notion de liaison chimique subit un nouveau changement. En 1939, le chimiste américain Linus Pauling présenta un concept de la liaison chimique fondé sur la mécanique quantique dans son livre, *La nature de la liaison chimique*. Sa théorie résolut enfin le paradoxe de la stabilité de la molécule du benzène.

Pauling a décrit les électrons qui forment la liaison chimique comme *résonnant* entre les atomes qu'ils joignent. Il a montré que dans certaines conditions, il est nécessaire de visualiser un électron comme occupant une position quelconque (avec différentes probabilités) parmi toutes celles possibles. On peut donc mieux représenter l'électron, et ses propriétés ondulatoires, comme étant distribué en une espèce de nuage, représentant la moyenne pondérée de la probabilité de ses positions particulières. Plus la distribution de l'électron est répartie d'une façon uniforme, plus le composé est stable. Une telle *stabilisation* par la résonance a plus de chances de se produire quand la molécule possède des liaisons conjuguées dans un plan et quand l'existence de symétries permet à l'électron, vu comme une particule, d'occuper des positions alternatives. Le cycle du benzène est plan et symétrique, et Pauling a montré que les liaisons de ce cycle n'étaient pas réellement doubles et simples en alternance, mais que les électrons étaient, en quelque sorte, distribués en une densité uniforme dont le résultat est que toutes les liaisons sont équivalentes, et dans l'ensemble plus fortes et moins réactives que les liaisons simples ordinaires.

Les structures résonnantes sont difficiles à représenter simplement sur une feuille de papier, bien qu'elles expliquent d'une façon satisfaisante les comportements chimiques. Les vieilles structures de Kekule, n'étant qu'une représentation approximative de la situation électronique réelle, sont et seront toujours utilisées.

La synthèse organique

Kolbe ayant réussi la synthèse de l'acide acétique, un autre chimiste, vers 1850, décida de synthétiser systématiquement et méthodiquement des substances organiques en laboratoire. C'était le Français Pierre Eugène

Marcelin Berthelot. Il prépara de nombreux composés organiques simples à partir de composés organiques encore plus simples tels que l'oxyde de carbone. Il les construisit de complexité croissante jusqu'à atteindre, entre autres choses, l'alcool éthylique. C'était de l'*alcool éthylique synthétique,* bien sûr, mais absolument *indiscernable* du produit naturel, parce que *c'était* le produit naturel.

L'alcool éthylique est un composé organique familier à tous et hautement apprécié. Le chimiste pouvait désormais fabriquer de l'alcool à partir du charbon, de l'air et de l'eau (le charbon fournit le carbone, l'air l'oxygène, et l'eau l'hydrogène), sans qu'il soit nécessaire de partir de fruits ou de grains ; et il se dotait par là même d'une nouvelle réputation de faiseur de miracles. De toute façon, cette découverte signa l'acte de naissance de la synthèse organique.

Cependant, pour les chimistes, Berthelot réalisa quelque chose de plus significatif encore. Il commença à fabriquer des produits qui n'existaient pas dans la nature. Il prit le *glycérol,* un composé découvert par Scheele en 1778 et obtenu par décomposition des graisses des organismes vivants, et le combina avec des acides que l'on ne trouve pas normalement dans les graisses naturelles (bien qu'on les trouve ailleurs dans la nature). Il obtint, de cette façon, des matières grasses qui n'étaient pas tout à fait celles trouvées dans l'organisme.

Berthelot fonda une nouvelle sorte de chimie organique – la synthèse de molécules que la nature ne fournit pas. Cela signifiait que, non seulement on pouvait faire des produits *synthétiques* qui soient des substituts de composés naturels – éventuellement inférieurs –, souvent difficiles ou impossibles à obtenir en quantités suffisantes, mais aussi des produits *synthétiques* qui soient supérieurs aux produits naturels.

Cette nouvelle perspective – améliorer la nature, plutôt que de simplement la compléter – s'est développée de façon colossale depuis que Berthelot en a ouvert la voie. Les premiers fruits furent obtenus dans le domaine des colorants.

LA PREMIÈRE SYNTHÈSE

La chimie organique est née en Allemagne. Wöhler et Liebig étaient tous deux allemands et furent suivis par d'autres, également très compétents. Il n'y eut pas de chimiste organicien en Angleterre jusqu'au milieu du XIXe siècle comparable aux chimistes allemands. En fait, les Anglais avaient une si piètre opinion de la chimie qu'ils n'enseignaient cette matière qu'à l'heure du déjeuner, n'espérant pas, ou peut-être ne désirant pas, que trop d'étudiants s'y intéressent. Il semble donc étonnant que la première synthèse ayant eu une répercussion mondiale ait été effectivement menée à bien en Angleterre.

Cet exploit fut réalisé de la façon suivante : en 1845, quand le Royal College of Science de Londres se décida enfin à programmer un enseignement de chimie de bon niveau, il fit appel au professeur August Wilhelm von Hofmann, un Allemand âgé de vingt-sept ans. Il fut engagé sur la suggestion du mari de la reine Victoria, le prince consort Albert (lui-même d'origine allemande).

Entre autres choses, Hofmann s'était intéressé aux composés du goudron, sur lesquels il avait travaillé sous la direction de Liebig. Le goudron est

un matériau noir et visqueux obtenu à partir du charbon quand il est fortement chauffé en l'absence d'air. Le goudron n'est pas un matériau séduisant, mais c'est une source précieuse de produits chimiques organiques. On a pu, en 1840 par exemple, en extraire de grandes quantités de benzène relativement pur, ainsi qu'un produit qui lui est apparenté et qui contient de l'azote : *l'aniline,* qu'Hofmann a été le premier à isoler.

Environ dix ans après son arrivée en Angleterre, Hofmann rencontra un jeune homme de dix-sept ans, étudiant la chimie au collège, William Henry Perkin. Hofmann avait un très bon flair pour détecter les talents et reconnaître le génie ; il engagea le jeune homme comme assistant et le fit travailler sur les composés du goudron. Perkin était d'une ardeur infatigable. Il installa un laboratoire chez lui et y travailla autant qu'au collège.

Un jour de 1856, Hofmann, s'intéressant aussi aux applications médicales de la chimie, rêva de la possibilité de synthétiser la quinine, une substance naturelle utilisée dans le traitement de la malaria. A l'époque, la représentation par les formules développées n'avait pas été élaborée. La seule chose connue de la quinine était sa composition atomique ; personne n'avait alors la moindre idée de la complexité de sa structure, qui n'a été correctement déterminée qu'en 1908.

Totalement ignorant de la complexité de la quinine, Perkin, âgé de dix-huit ans, s'attaqua en toute innocence à sa synthèse. Il partit de l'allyle toluidine, un des composés du goudron. Cette molécule était composée de la moitié environ de chacun des types d'atomes de la molécule de quinine. Perkin pensait qu'il pourrait obtenir une molécule de quinine en additionnant deux molécules d'allyle toluidine et en leur ajoutant les atomes d'oxygène manquants (en y mélangeant du bichromate de potassium, par exemple, ce qui est un moyen d'ajouter des atomes d'oxygène à des composés chimiques).

Naturellement, ce premier essai de Perkin n'aboutit à rien, ou plutôt à une soupe sale et brun-rouge. Il essaya ensuite l'aniline au lieu de l'allyle toluidine et obtint un brouet noirâtre. Cependant, cette fois, il lui sembla y voir des reflets pourpres. Il ajouta de l'alcool à ce mélange infâme, et le liquide prit une merveilleuse couleur pourpre. Perkin pensa alors qu'il venait de découvrir quelque chose qui pouvait être utilisé comme colorant.

Les colorants ont toujours été des substances très recherchées et chères. Il n'existait que peu de bons colorants – des colorants qui teignent les tissus de manière permanente et éclatante, qui ne passent ni ne déteignent. Il y avait un bleu indigo foncé, obtenu à partir de la plante indigo et de sa proche parente, le *pastel.* Il y avait la *pourpre de Tyr,* extraite d'un escargot (appelée ainsi parce que les anciens habitants de Tyr s'enrichirent de sa fabrication ; à la fin de l'Empire romain, les enfants royaux naissaient dans une pièce tapissée de pourpre tyrienne, d'où l'expression « né dans la pourpre »). Enfin, il y avait le rouge *alizarine,* extrait de la garance (alizarine provient du mot arabe signifiant « le jus »). A ces héritages des temps anciens et du Moyen Age, les teinturiers ajoutèrent plus tard des colorants provenant des tropiques ainsi que des pigments minéraux (principalement utilisés de nos jours dans les peintures).

D'où l'exaltation de Perkin à la pensée que cette substance pourpre pouvait être un colorant. Suivant le conseil d'un ami, il en envoya un échantillon à un teinturier écossais. La réponse revint très rapidement,

disant que le composé pourpre avait de bonnes propriétés colorantes. Pourrait-il être fourni à bas prix ? Perkin fit breveter le colorant (on discuta fermement pour savoir si un garçon de dix-huit ans pouvait déposer un brevet, qu'il obtint finalement), arrêta ses études et entra dans les affaires.

Le projet de Perkin n'était pas facile. Il avait à concevoir l'équipement et à préparer ses propres matières premières à partir du goudron ; tout était à faire. Il réussit cependant à produire ce qu'il appela de l'*aniline pourpre* en moins de six mois ; un composé que l'on ne trouve pas dans la nature, et supérieur à tous les colorants naturels pour ce type de couleur.

Les teinturiers français, qui adoptèrent le nouveau colorant plus vite que ne le firent les Anglais, plus conservateurs, appelèrent cette couleur *mauve,* de la plante du même nom (en latin *malva*) ; et le colorant fut connu sous le nom de *mauvéine.* Il fit rapidement fureur, et Perkin s'enrichit. A vingt-trois ans, il faisait autorité dans le monde en matière de colorants.

La porte était ouverte. Inspirés par le succès étonnant de Perkin, de nombreux chimistes organiciens se lancèrent dans la synthèse des colorants et un certain nombre y réussit. Hofmann lui-même se convertit à cette nouvelle spécialité et synthétisa en 1858 un colorant rouge pourpre, auquel les teinturiers français (alors les arbitres de la mode dans le monde) donnèrent le nom de *magenta,* nom de la ville italienne où les Français battirent les Autrichiens en 1859.

De retour en Allemagne en 1865, Hofmann, tout à son nouvel intérêt pour les colorants, découvrit un nouveau groupe de colorants violets, encore connus sous le nom de *violets de Hofmann.* Il y avait au moins 3 500 colorants commercialisés, à la moitié du XXe siècle.

Les chimistes ont aussi synthétisé des colorants naturels en laboratoire. L'Allemand Karl Graebe et Perkin synthétisèrent l'alizarine en 1869 (Graebe déposa son brevet un jour avant Perkin). Le chimiste allemand Adolf von Baeyer mit au point en 1880 une technique pour synthétiser l'indigo. (Von Baeyer reçut le prix Nobel de chimie en 1905 pour son travail sur les colorants.)

Perkin se retira des affaires en 1874, à trente-cinq ans, et revint à ses premières amours : la recherche. Dès 1875, il avait réussi à synthétiser la *coumarine* (une substance naturelle qui a la bonne odeur du foin coupé). Il créa ainsi l'industrie du parfum synthétique.

Perkin tout seul ne pouvait pas maintenir la suprématie anglaise face au grand essor de la chimie organique allemande. A la fin du siècle, les « synthétiques » devinrent presque un monopole allemand. Ce fut le chimiste allemand Otto Wallach qui mena à bien le travail sur les parfums synthétiques que Perkin avait amorcé. Wallach reçut le prix Nobel de chimie pour ses recherches en 1910. Le chimiste croate Leopold Ruzicka, enseignant en Suisse, fut le premier à synthétiser le *musc,* un composé important en parfumerie. Il partagea le prix Nobel de chimie en 1938. Cependant, lors de la Première Guerre mondiale, la Grande-Bretagne et les États-Unis furent coupés des produits des laboratoires chimiques allemands et durent développer leur propre industrie chimique.

LES ALCALOÏDES ET LES CALMANTS DE LA DOULEUR

La chimie organique n'aurait progressé qu'à petits pas si les chimistes s'étaient contentés de découvertes fortuites, telles que celles de Perkin. Heureusement, les formules développées de Kekule, introduites trois ans après la découverte de Perkin, permirent de dessiner des modèles de la molécule organique. Les chimistes ne se contentaient plus de préparer la quinine au juger, mais disposaient de méthodes permettant d'atteindre, étape par étape, les sommets structuraux des molécules par des chemins qu'ils savaient déterminer à l'avance.

Les chimistes apprirent à modifier un groupement d'atomes, ouvrir les cycles d'atomes ou former des cycles à partir de chaînes linéaires, couper des groupements d'atomes en deux ou ajouter un à un des atomes de carbone à une chaîne. On se réfère encore souvent à une méthode spécifique, modifiant la structure d'une molécule organique d'une façon donnée, par le nom du chimiste qui l'a décrite en premier. Par exemple, Perkin a découvert une méthode pour ajouter un groupement de deux atomes de carbone à un autre composé en chauffant certaines substances en présence d'*anhydride acétique* et d'*acétate de sodium* ; et la méthode est toujours appelée la *réaction de Perkin.* Le professeur de Perkin, Hofmann, découvrit qu'un cycle d'atomes comprenant un atome d'azote pouvait être traité par l'*iodure de méthyle* en présence d'un sel d'argent, de telle façon que le cycle est finalement ouvert et l'atome d'azote est éliminé. C'est la *dégradation d'Hofmann.* En 1877, le chimiste français Charles Friedel, travaillant avec le chimiste américain James Mason Crafts, mit au point une façon de lier une courte chaîne carbonée à un cycle benzénique en les chauffant en présence de chlorure d'aluminium. Cette réaction est connue sous le nom de réaction de *Friedel-Crafts.*

En 1900, le chimiste français Victor Grignard découvrit que le magnésium, sous forme métallique, convenablement utilisé, permettait de fixer une grande variété de chaînes carbonées de manières différentes. Il présenta ce travail dans sa thèse de doctorat d'État. Le développement de ces *réactions de Grignard* lui valut de partager le prix Nobel en 1912. Le chimiste français Paul Sabatier, qui partagea ce prix avec lui, avait découvert (avec Jean-Baptiste Senderens) une méthode utilisant le nickel finement divisé pour provoquer l'addition d'atomes d'hydrogène dans des doubles liaisons de chaînes carbonées. C'est la réduction de *Sabatier-Senderens.*

En 1928, les chimistes allemands Otto Diels et Kurt Alder découvrirent une méthode pour lier les deux extrémités d'une chaîne de carbone aux extrémités opposées d'une double liaison d'une autre chaîne de carbone, formant ainsi un cycle d'atomes. Pour la découverte de cette *réaction de Diels-Alder,* ils partagèrent le prix Nobel de chimie en 1950.

En d'autres termes, en notant sur les formules développées les changements produits lorsqu'une substance est soumise à des traitements chimiques variés, les chimistes organiciens établirent peu à peu une série de règles de base indiquant comment changer à volonté un composé en un autre. Cela n'a pas été facile. Chaque composé, chaque changement présentait ses propres particularités et difficultés. Les grandes avenues étaient cependant tracées et les chimistes organiciens ayant du talent y

trouvèrent des signes évidents de progrès dans ce qui auparavant avait semblé une jungle.

Savoir comment un groupe donné d'atomes se comporte peut aussi servir à déterminer la structure de composés inconnus. Par exemple, quand des alcools simples réagissent avec du sodium métallique et libèrent de l'hydrogène, seul l'hydrogène lié à l'atome d'oxygène est libéré, les hydrogènes liés aux atomes de carbone ne le sont pas. D'autre part, certains composés organiques captent des atomes d'hydrogène dans des conditions appropriées, et d'autres non. Il s'avère que les composés qui captent de l'hydrogène possèdent généralement des liaisons doubles ou triples et que l'hydrogène s'additionne à ces liaisons. A partir de ce genre d'information, une analyse chimique des composés organiques entièrement nouvelle est née. On détermine la nature des groupements d'atomes, plutôt que simplement le nombre et la variété des différents atomes présents. La libération d'hydrogène en présence de sodium indique la présence d'hydrogène lié à l'oxygène dans le composé. La fixation d'hydrogène indique la présence d'une liaison double ou triple. Si la molécule est trop compliquée pour être analysée dans son entier, elle peut être cassée en des entités plus simples par des méthodes bien définies. On peut alors déterminer les structures de ces morceaux de molécules plus petits et reconstituer la formule de la molécule d'origine déduite de ceux-ci.

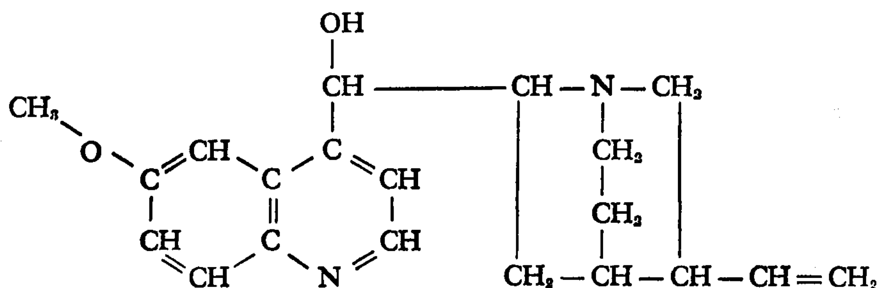
Les chimistes ont déterminé la structure de certains composés organiques naturels (analyse), en se servant des formules développées comme outils et comme guides. Ils pouvaient alors les reproduire soit exactement, soit approximativement en laboratoire (synthèse). Un produit rare et cher, ou difficile à se procurer dans la nature, pouvait de ce fait devenir disponible à bon marché et en quantité en laboratoire. Ou, comme c'est le cas des colorants dérivés du goudron, on pouvait créer en laboratoire quelque chose qui apportait une réelle amélioration par rapport aux substances existant dans la nature.

Un cas significatif d'amélioration voulue d'un produit naturel est celui de la cocaïne, trouvée dans les feuilles du coca, arbrisseau originaire de Bolivie et du Pérou, mais qui est maintenant cultivé principalement à Java. La cocaïne est un exemple d'*alcaloïde*, comme la strychnine, la morphine et la quinine, tous déjà mentionnés. C'est un produit végétal contenant de l'azote qui, à faible dose, a des effets physiologiques importants sur l'homme. Selon la dose, les alcaloïdes peuvent soigner ou tuer. La mort la plus célèbre causée par un alcaloïde est celle de Socrate, tué par la *cicutine*, l'alcaloïde de la ciguë.

La structure moléculaire des alcaloïdes est, dans certains cas, extraordinairement compliquée, mais cela n'a fait qu'aiguiser la curiosité des chimistes. Le chimiste anglais Robert Robinson s'est attaqué systématiquement aux alcaloïdes. Il a déterminé en 1925 la structure de la morphine (à l'exception d'un atome douteux), et celle de la strychnine en 1946. Il a reçu le prix Nobel de chimie en 1947.

Robinson avait simplement déterminé la structure des alcaloïdes, sans s'en servir comme guide pour leur synthèse éventuelle. Ce fut le chimiste américain Robert Burns Woodward qui synthétisa la quinine en 1944 avec son collègue américain William von Eggers Doering. On se rappelle la quête désordonnée de Perkin pour synthétiser ce produit qui avait donné des

résultats si spectaculaires. Si vous êtes curieux, voici la structure de la quinine :



Que Woodward et von Doering aient résolu le problème n'est pas seulement dû à leur talent. Ils avaient aussi à leur disposition les nouvelles théories électroniques de la structure et du comportement moléculaire développées par des hommes tels que Pauling. Woodward continua à synthétiser différentes molécules complexes qui auraient, avant cette époque, représenté des défis insurmontables. Par exemple, en 1954, il synthétisa la strychnine.

Bien avant que la structure des alcaloïdes soit connue, certains d'entre eux – la cocaïne en particulier – étaient cependant l'objet d'un grand intérêt de la part des médecins. On savait que les Indiens d'Amérique du Sud mâchaient de la feuille de coca, et qu'ils y trouvaient un antidote à la fatigue et une source de sensations bienfaisantes. Le médecin écossais Robert Christison a introduit la plante en Europe. Ce n'est pas le seul cadeau fait à la médecine par les sorciers et les sorcières des sociétés préscientifiques. Il y a aussi la quinine et la strychnine, ainsi que l'opium, la digitale, le curare, l'atropine, la strophantidine et la réserpine. De plus, le tabac, les noix de bétel que l'on mâche, l'alcool, les drogues comme la marijuana et le peyotl sont tous des héritages des sociétés primitives.

La cocaïne n'était pas seulement un euphorisant. Les docteurs découvrirent qu'elle insensibilisait le corps aux sensations de douleur temporairement et localement. En 1884, le médecin américain Carl Koller remarqua que la cocaïne pouvait agir comme calmant de la douleur quand on en mettait sur les muqueuses autour de l'œil. Les opérations de l'œil pouvaient être faites de manière indolore. Les dentistes utilisaient aussi la cocaïne pour arracher les dents sans douleur.

Cette propriété intéressa considérablement les médecins, car au XIXe siècle une des grandes victoires médicales fut celle remportée sur la douleur. En 1799, Humphry Davy avait préparé de l'*oxyde nitreux* (N_2O), un gaz, et en avait étudié les effets. Il trouva que lorsqu'il était inhalé, il abolissait les inhibitions, de sorte que toute personne qui le respirait se mettait à rire, pleurer, ou agir de façon stupide, d'où son nom de *gaz hilarant.*

Au début des années 1840, un chercheur américain, Gardner Quincy Cotton, découvrit que l'oxyde nitreux calmait les sensations douloureuses, et en 1844, un dentiste américain, Horace Wells, l'utilisa en chirurgie dentaire. Mais un produit meilleur était déjà apparu.

En 1842, le chirurgien américain Crawford Williamson Long avait utilisé de l'éther pour endormir un patient alors qu'il lui arrachait une dent. En

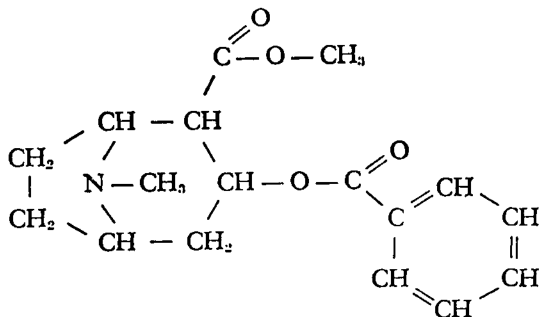
1846, le dentiste américain William Thomas Green Morton fit une opération chirurgicale sous éther au Massachusetts General Hospital. On lui attribue généralement le crédit de cette découverte, car Long n'a décrit son exploit dans un journal médical qu'après la démonstration publique de Morton, et les expérimentations plus précoces de Wells avec l'oxyde nitreux sont passées inaperçues.

Le médecin et poète américain Oliver Wendell Holmes suggéra que l'on appelle les composés calmant la douleur *anesthésiques* (du mot grec signifiant « sans sensation »). Certaines personnes à cette époque estimaient que les anesthésiques étaient une entreprise sacrilège cherchant à éviter la douleur que Dieu inflige aux êtres humains. L'utilisation par le médecin écossais James Young Simpson de l'anesthésie au cours de l'accouchement de la reine Victoria lui donna ses lettres de noblesse.

La chirurgie, qui était une boucherie rappelant les chambres de torture, devenait plus humaine grâce à l'anesthésie et à certaines conditions antiseptiques. Elle parvenait même à sauver des vies. C'est pour cette raison que toute découverte dans le domaine de l'anesthésie était suivie avec grand intérêt. La cocaïne, un *anesthésique local,* calmant la douleur dans une aire limitée sans provoquer d'état d'inconscience ni de perte de sensations (comme dans le cas d'*anesthésies générales* par l'éther), présentait un intérêt particulier.

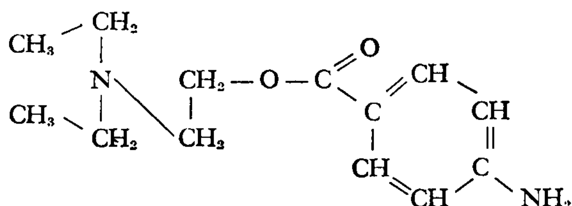
Il existe cependant plusieurs inconvénients liés à la cocaïne. Elle peut d'abord induire des effets secondaires très nocifs, allant jusqu'à tuer certains malades. Ensuite, elle peut entraîner une accoutumance et doit être utilisée avec parcimonie et prudence. La cocaïne est l'une de ces *drogues* dangereuses qui ne diminuent pas seulement la douleur mais aussi les autres sensations désagréables, et donnent à l'utilisateur l'illusion de l'euphorie. Ce dernier s'habitue à la drogue de sorte qu'il a besoin de doses de plus en plus fortes et, malgré les effets néfastes de la drogue sur son corps, il devient si dépendant des illusions qu'elle apporte qu'il ne peut pas s'arrêter d'en prendre sans développer les symptômes pénibles du *manque.* Une telle *toxicomanie* vis-à-vis de la cocaïne et d'autres drogues similaires est un problème social important. On produit illégalement jusqu'à vingt tonnes de cocaïne chaque année, qui sont vendues avec d'énormes bénéfices pour quelques-uns, une énorme misère pour beaucoup et ce, malgré les efforts internationaux pour arrêter ce trafic. Enfin, la molécule de cocaïne est fragile, et lorsqu'on la chauffe pour la stériliser de toute contamination bactérienne, on peut altérer la molécule, ce qui modifie ses effets anesthésiques.

La structure de la molécule de la cocaïne est complexe :



Le double cycle à gauche est la partie fragile de la molécule, c'est aussi la partie difficile à synthétiser. La synthèse de la cocaïne ne fut obtenue qu'en 1923 par Richard Willstätter. Cependant, les chimistes imaginèrent qu'ils pouvaient synthétiser des composés similaires, où le double cycle ne serait pas fermé, et construire plus facilement des molécules plus stables. Les substances synthétiques devraient avoir les propriétés anesthésiques de la cocaïne sans en avoir les effets secondaires.

Pendant environ vingt ans les chimistes allemands se sont attelés au problème, produisant des douzaines de composés, dont certains étaient assez bons. La modification la plus heureuse fut obtenue en 1909, avec le composé suivant :



Si vous comparez cette formule avec celle de la cocaïne, vous verrez la similitude, mais aussi le fait très important que le double cycle n'existe plus. Cette molécule plus simple, stable, facile à synthétiser, ayant de bonnes propriétés anesthésiques et peu d'effets secondaires, n'existe pas dans la nature. C'est une *imitation synthétique,* meilleure que le produit naturel. Cette molécule s'appelle procaïne, plus connue du public sous la marque Novocaïne.

Le calmant de la douleur probablement le plus connu et le plus efficace est la *morphine*. Son nom même vient du mot grec « dormir ». C'est un dérivé purifié de la sève de l'opium ou *laudanum,* utilisé depuis des siècles par les peuples pour combattre la douleur et la tension quotidienne. C'est un merveilleux cadeau pour la personne qui souffre, mais elle fait aussi courir le danger mortel de la toxicomanie. Les efforts mis en œuvre pour trouver un succédané se sont avérés à double tranchant. On créa en 1898 un dérivé synthétique, le *diacétyle morphine,* plus connu sous le nom d'*héroïne,* en pensant qu'il serait moins dangereux. Au contraire, la drogue se révéla être la plus dangereuse de toutes.

Les *sédatifs* (inducteurs de sommeil) les moins dangereux sont l'*hydrate de chloral* et les *barbituriques*. Le premier des barbituriques fut découvert en 1902, et ce sont aujourd'hui les composants habituels les plus communs des *somnifères*. Relativement sans danger quand ils sont utilisés correctement, ils peuvent malgré tout provoquer une accoutumance et, pris en excès, la mort. Le recours à une dose excessive de barbiturique est une méthode assez populaire de suicide ou de tentative de suicide, parce que la mort vient doucement, au terme d'un sommeil de plus en plus profond.

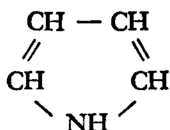
Le sédatif le plus courant, celui utilisé depuis la nuit des temps, est naturellement l'alcool. La fermentation des jus de fruits ou des grains était connue à l'époque préhistorique. Pour produire des liqueurs plus

fortes que ce que l'on pouvait obtenir naturellement, la distillation fut introduite au Moyen Age. Dans les régions du globe où l'alimentation en eau entraîne presque immanquablement le risque de fièvre typhoïde ou de choléra, il est préférable de consommer modérément des vins légers, ce qui est accepté par la société. Cela rend difficile de traiter l'alcool comme une drogue, ce qu'il est pourtant, car il provoque une accoutumance au même titre que la morphine et, utilisé en grande quantité, il fait beaucoup plus de dégâts. L'interdiction légale par la loi de la vente d'alcool ne semble d'aucune aide. Il est certain que l'expérience de la Prohibition (1920-1933) aux États-Unis a été un échec désastreux. L'alcoolisme est néanmoins de plus en plus traité comme une maladie – et c'en est une – que comme un signe de déchéance morale. Les symptômes aigus de l'alcoolisme (*delirium tremens*) ne sont probablement pas tant dus à l'alcool lui-même qu'à la carence en vitamines qui apparaît chez ceux qui boivent beaucoup et mangent peu.

LES PROTOPORPHYRINES

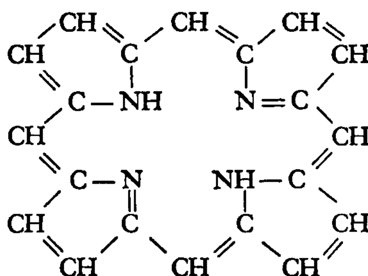
L'homme a maintenant à sa disposition toutes sortes de produits synthétiques aux multiples possibilités qu'il peut utiliser en bien ou en mal : explosifs, gaz empoisonnés, insecticides, désherbants, antiseptiques, désinfectants, détergents, drogues... la liste est sans fin. La synthèse n'est pas seulement au service des besoins du consommateur, elle est aussi au service de la recherche pure en chimie. Un composé complexe, provenant des tissus vivants, et dont toutes les réactions ont été étudiées, ne peut recevoir parfois qu'une structure moléculaire provisoire. Dans ce cas-là, une façon de s'en sortir est de synthétiser le composé au moyen de réactions connues, afin d'obtenir une structure moléculaire semblable à celle qui a été proposée. Si les propriétés de la molécule obtenue sont identiques à celles du produit sujet de la recherche, un chimiste peut être sûr que la structure proposée est la bonne.

Un cas typique et impressionnant est celui des *hémoglobines*, le composant principal des globules rouges et le pigment qui donne sa couleur rouge au sang. Le chimiste français L.R. Le Canu a séparé en 1831 l'hémoglobine en deux parties, dont la plus petite, appelée l'*hème*, représente 4 % de la masse de l'hémoglobine. Il a été trouvé pour l'hème la formule brute $C_{34}H_{32}O_4N_4Fe$. Étant donné que l'hème était connu pour exister dans d'autres substances d'importance vitale, dans le règne végétal et le règne animal, la structure de cette molécule a pris une grande importance auprès des biochimistes. Pendant près d'un siècle après la découverte de Le Canu, tout ce que l'on a su faire était de casser l'hème en molécules plus petites. L'atome de fer (Fe) a été facilement arraché, et ce qui restait de la molécule se cassait alors en morceaux représentant environ un quart de celle-ci. Ces fragments étaient des *pyrroles*, des cycles de cinq atomes : quatre carbones et un azote. Le pyrrole lui-même a la structure suivante :



Les pyrroles obtenus à partir de l'hème possèdent des radicaux de un ou deux atomes de carbone attachés au cycle à la place d'un ou plusieurs atomes d'hydrogène.

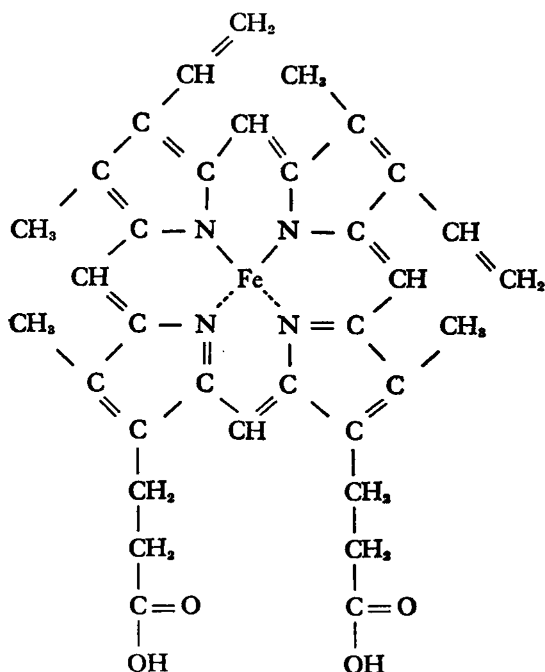
Dans les années vingt, le chimiste Hans Fischer fit de nouveaux progrès. Les pyrroles représentant un quart de la taille de l'hème original, il décida d'assembler les quatre pyrroles pour voir ce qu'il obtiendrait. Il obtint finalement un composé à quatre cycles qu'il appela *porphine* (du grec signifiant « pourpre », parce qu'il était de couleur pourpre). La porphine devrait être ainsi :



Mais comme les pyrroles obtenus à partir de l'hème contiennent des petites *chaînes latérales* attachées aux cycles, celles-ci demeurent inchangées quand les pyrroles sont assemblés pour former la porphine. Les porphines ayant différentes chaînes latérales constituent une famille de composés appelés *porphyrines*. Les composés ayant la chaîne latérale particulière trouvée dans l'hème sont les *protoporphyrines*. Après avoir comparé les propriétés de l'hème avec celles des porphyrines qu'il avait synthétisées, Fischer en conclut que l'hème (sans atome de fer) était une protoporphyrine. Mais laquelle ? On pouvait former au moins quinze protoporphyrines distinctes (chacune ayant une disposition différente des chaînes latérales) à partir des pyrroles variés obtenus de l'hème. Selon le raisonnement de Fischer, n'importe laquelle de ces quinze molécules ainsi obtenues pouvait être l'hème lui-même.

Une réponse directe pouvait être obtenue en synthétisant ces quinze molécules et en testant les propriétés de chacune d'elles. Fischer et ses étudiants préparèrent les quinze composés à l'aide de réactions chimiques très compliquées, qui permettaient d'aboutir à une structure unique. Au fur et à mesure de la synthèse des différentes protoporphyrines, ils comparaient leurs propriétés avec celles de la protoporphyrine naturelle de l'hème.

En 1928, Fischer découvrit que la neuvième protoporphyrine dans sa série était celle qu'il cherchait. Depuis lors, la protoporphyrine naturelle s'appelle *protoporphyrine IX*. C'était une opération simple que de convertir la protoporphyrine IX en hème par l'addition de fer. Les chimistes étaient enfin certains de connaître la structure de cet important composé. Voici la structure de l'hème, telle qu'elle a été établie par Fischer :



Pour son exploit, Fischer reçut le prix Nobel de chimie en 1930.

LES NOUVEAUX PROCÉDÉS

Tous les succès remportés par la synthèse organique chimique pendant le XIXe siècle et la première moitié du XXe, aussi grands fussent-ils, ont été obtenus par les moyens des alchimistes des temps anciens : mélanger et chauffer les substances. La chaleur était une façon sûre de fournir de l'énergie aux molécules et de les faire interagir, mais les interactions se font habituellement au hasard et ont lieu grâce à des intermédiaires instables, qui existent brièvement, dont la nature ne peut qu'être devinée.

Les chimistes avaient besoin d'une méthode directe, plus raffinée, pour produire des molécules dans un état d'énergie élevée, une méthode qui produirait un groupe de molécules se déplaçant toutes à la même vitesse et dans la même direction. Cette méthode éliminerait la nature aléatoire des interactions, car quel que soit le comportement d'une molécule, toutes réagiraient de la même manière. Une méthode consistait à accélérer les ions dans un champ électrique, un peu comme les particules subatomiques sont accélérées dans un cyclotron.

En 1964, le chimiste américain d'origine allemande, Richard Leopold Wolfgang, accéléra des ions et des molécules à de hautes énergies, au moyen de ce qu'on pourrait appeler un *accélérateur chimique*, produisant des vitesses ioniques que la chaleur ne permet d'atteindre qu'à des températures de 10 000 °C à 100 000 °C. De plus, tous les ions se déplacent dans la même direction. Si on fournit des électrons à des ions ainsi accélérés, ils s'y aggripent, formant alors des molécules neutres qui se

déplaceront encore à très grande vitesse. De tels faisceaux neutres ont été produits en 1969 par le chimiste américain Leonard Wharton.

Les ordinateurs peuvent aider à appréhender les étapes intermédiaires et brèves des réactions chimiques. Cela demandait la mise au point des équations quantiques chimiques gouvernant l'état des électrons dans différentes combinaisons d'atomes, et la définition des événements qui pouvaient avoir lieu lors des collisions. En 1968 par exemple, un ordinateur dirigé par le chimiste américain d'origine italienne Enrico Clementi contrôla, en circuit fermé de télévision, la *collision* de l'ammoniac et de l'acide chlorhydrique pour former le chlorure d'ammonium, l'ordinateur déterminant les événements qui devaient avoir lieu. L'ordinateur indiquait que le chlorure d'ammonium formé pouvait exister sous forme de gaz à haute pression et à la température de 700 °C. On ignorait cette possibilité, mais elle fut démontrée expérimentalement quelques mois plus tard.

Durant les dernières décades, les chimistes ont créé des outils tout à fait nouveaux, à la fois théoriques et expérimentaux. Des détails subtils des réactions, qui n'ont pas été disponibles jusqu'à présent, seront bientôt connus, et de nouveaux produits – impossibles à obtenir auparavant, ou qu'on ne pouvait obtenir qu'en petits échantillons – seront formés. Nous sommes peut-être au seuil de merveilles inattendues.

Polymères et plastiques

Quand nous considérons des molécules telles que l'hème et la quinine, nous nous approchons d'un niveau de complexité auquel même les chimistes modernes ne peuvent faire face que difficilement. La synthèse de tels composés requiert des étapes si nombreuses et une telle variété de procédés que l'on ne peut guère s'attendre à en produire en quantité sans l'aide d'un organisme vivant (autre que le chimiste). Il n'y a cependant pas de quoi en faire un complexe d'infériorité. Le tissu vivant lui-même, à ce niveau de complexité, approche les limites de ses capacités. Peu de molécules dans la nature sont aussi compliquées que l'hème et la quinine.

Bien sûr, il y a des substances naturelles composées de centaines de milliers, voire de millions d'atomes, mais ce ne sont pas vraiment des molécules individuelles, c'est-à-dire construites d'une seule pièce. Ces grandes molécules sont plutôt construites d'unités liées entre elles, comme les perles d'un collier. En général, le tissu vivant synthétise quelques petits composés assez simples et les réunit ensuite en chaînes. Et cela, comme nous le verrons, le chimiste sait aussi le faire.

CONDENSATION ET GLUCOSE

Dans le tissu vivant, cette union de petites molécules (*condensation*) s'accompagne habituellement de l'élimination de deux atomes d'hydrogène et d'un atome d'oxygène (qui se combinent pour former une molécule d'eau) à chaque point de raccordement. Le processus peut être inversé (dans le tube à essai et dans le corps) par l'addition d'eau, et les maillons de la chaîne peuvent être desserrés et séparés. Cette inversion de la

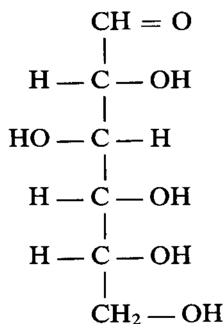
condensation s'appelle l'*hydrolyse*, du grec signifiant « décomposition par l'eau ». Dans le tube à essai, l'hydrolyse de ces longues chaînes peut être accélérée par de nombreuses méthodes, la plus courante étant l'addition d'une quantité donnée d'acide au mélange.

Les premières recherches sur la structure chimique des grandes molécules remontent à 1812, quand le chimiste russe Gottlieb Sigismund Kirchhoff découvrit qu'en faisant bouillir de l'amidon avec de l'acide, on produisait un sucre ayant des propriétés identiques à celles du glucose (le sucre obtenu à partir du raisin). En 1819, le chimiste français Henri Braconnot obtint aussi du glucose en faisant bouillir des produits végétaux variés tels que la sciure de bois, le lin, l'écorce, qui contiennent tous un composé appelé la cellulose. Il était facile de deviner que l'amidon et la cellulose étaient formés d'unités de glucose, mais pour connaître les détails de la structure moléculaire de l'amidon et de la cellulose, on dut attendre que la structure du glucose soit connue. Avant le temps des formules développées, tout ce que l'on savait alors du glucose était sa formule brute : $C_6H_{12}O_6$, de laquelle on pouvait déduire qu'il y avait une molécule d'eau, H_2O, attachée à chacun des six atomes de carbone ; d'où l'appellation d'*hydrates de carbone* (« les carbones aqueux ») attribuée au glucose et aux composés similaires.

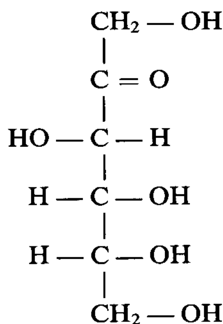
La formule développée du glucose a été établie en 1886 par le chimiste allemand Heinrich Kiliani. Il démontra que cette molécule est constituée d'une chaîne de six atomes de carbone auxquels sont attachés des atomes d'hydrogène et des groupements d'oxygène-hydrogène. Il n'y a pas, en fait, d'association de molécule d'eau intacte sur un atome de carbone, dans la molécule de glucose.

A la fin du siècle, le chimiste allemand Emil Fischer a étudié le glucose en détail et élucidé l'arrangement des groupements oxygène-hydrogène autour des atomes de carbone, dont quatre étaient asymétriques. Il y a seize arrangements possibles de ces groupements, et donc seize isomères optiques, chacun avec ses propriétés. Les chimistes ont réellement synthétisé les seize isomères, dont seuls quelques-uns se trouvent dans la nature. C'est à partir des résultats de son travail sur l'activité optique de ces sucres que Fischer suggéra l'établissement des séries de composés L et D. Pour avoir établi la chimie des sucres sur des fondations structurales fermes, Fischer reçut, en 1902, le prix Nobel de chimie.

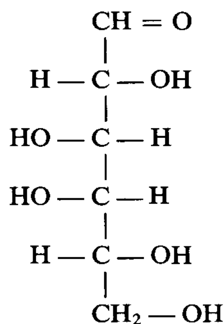
Les formules développées du glucose et de deux autres sucres fréquemment rencontrés, le fructose et le galactose, sont données ci-dessous :



Glucose



Fructose



Galactose

Ce sont les structures les plus simples qui représentent de façon adéquate les asymétries de la molécule ; mais en réalité, les molécules ont la forme d'un cycle plan, chaque cycle étant fait de cinq (quelquefois quatre) atomes de carbone et d'un atome d'oxygène.

Quand les chimistes eurent connu la structure des sucres simples, il leur fut relativement aisé de déterminer la façon dont ces sucres s'associaient en des composés plus complexes. Par exemple, une molécule de glucose et une molécule de fructose peuvent se condenser pour donner le saccharose – le sucre que nous utilisons à table. Le glucose et le galactose se combinent pour former le lactose que, dans la nature, on ne trouve que dans le lait.

Il n'y a pas de raison pour que les condensations ne continuent pas à l'infini, ce qui se passe dans l'amidon et la cellulose. Chacune de ces molécules consiste en de longues chaînes d'unités de glucose, condensées selon un schéma particulier.

L'ordre de ce schéma est important : bien que chacun des deux composés soit construit à partir de la même unité, ils sont pourtant profondément différents. L'amidon, sous une forme ou sous une autre, constitue la plus grande part du régime alimentaire de l'homme, tandis que la cellulose est parfaitement indigeste. La différence dans le schéma de condensation, résolu par le travail minutieux des chimistes, peut être représentée ainsi : supposons une molécule de glucose vue soit à l'endroit (symbolisée par *u*) soit à l'envers (symbolisée par *n*). La molécule d'amidon peut être visualisée comme étant composée d'une chaîne de molécules de glucose telle que « .....uuuuuuuuuu..... », tandis que la cellulose est faite de « .....unununununu..... ». Nos sucs digestifs peuvent hydrolyser la liaison « uu » de l'amidon, le décomposant en glucose, qui peut alors être absorbé pour obtenir de l'énergie. Les mêmes sucs digestifs sont incapables de couper les liaisons « un » ou « nu » de la cellulose, et toute la cellulose que nous ingérons traverse le tube digestif et en ressort.

Certains micro-organismes peuvent digérer la cellulose, mais aucun être supérieur ne le peut. Certains de ces micro-organismes vivent dans les intestins des ruminants ou des termites, par exemple. C'est grâce à eux que les vaches, pour notre avantage, peuvent vivre d'herbe et que les termites, souvent pour notre déconvenue, peuvent vivre de bois. Les micro-organismes forment du glucose en quantité à partir de la cellulose, ils utilisent ce dont ils ont besoin, et l'hôte utilise le surplus. Les micro-organismes fournissent la nourriture transformée, tandis que l'hôte fournit la matière première et le logement. Cette coopération entre deux formes de vie pour leur bénéfice commun est appelée la *symbiose,* du grec « vie avec ».

POLYMÈRES CRISTALLINS ET AMORPHES

Christophe Colomb découvrit les indigènes d'Amérique du Sud jouant avec des balles faites à partir de la sève durcie d'une plante. Christophe Colomb, ainsi que les autres explorateurs qui visitèrent l'Amérique du Sud durant les deux siècles qui suivirent, était fasciné par ces balles rebondissantes (obtenues à partir de la sève d'arbres du Brésil). Des échantillons ont finalement été rapportés en Europe comme objets de

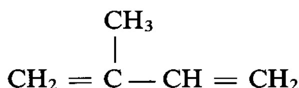
curiosité. Aux environs de 1770, Joseph Priestley (qui devait bientôt découvrir l'oxygène) remarqua qu'un morceau de ce matériau rebondissant gommait les marques de crayon ; il inventa le nom de *gomme caoutchouc.* Les Anglais l'appellent *gomme indienne (Indian rubber),* parce qu'elle provient des « Indes » (le nom donné à l'origine par Christophe Colomb au Nouveau Monde).

Par la suite, on a trouvé d'autres utilisations au caoutchouc. En 1823, un Écossais du nom de Charles Macintosh a déposé un brevet pour des vêtements de pluie, faits d'une couche de caoutchouc comprise entre deux couches de tissu ; depuis, les imperméables sont encore quelquefois appelés *Mackintosh* (avec un *k* en plus).

L'inconvénient, c'est que par temps chaud le caoutchouc devient collant, tandis que par temps froid il est dur et raide. De nombreuses personnes ont essayé de trouver un moyen de traiter le caoutchouc pour en éliminer ces caractéristiques indésirables ; parmi celles-ci, un Américain, Charles Goodyear, a travaillé dans ce sens, bien que n'étant pas chimiste. Un jour de 1839, il renversa accidentellement un mélange de caoutchouc et de soufre sur un fourneau brûlant. Il le racla rapidement et trouva, à son grand étonnement, que le mélange caoutchouc-soufre était sec tandis qu'il était encore chaud. Il avait un échantillon de caoutchouc qui ne devenait ni collant en le chauffant, ni raide après refroidissement. Il restait souple et élastique en toutes circonstances.

En ajoutant du soufre au caoutchouc, on obtient la *vulcanisation* (d'après Vulcain, le dieu du feu chez les Romains). Cette découverte de Goodyear a établi les bases de l'industrie du caoutchouc. (Goodyear n'a jamais reçu de récompense pour cette découverte qui rapporta pourtant des millions de dollars. Il passa sa vie à se battre pour les redevances de ses brevets et mourut couvert de dettes.)

La structure moléculaire du caoutchouc fut connue en 1879, quand le chimiste français Gustave Bouchardat chauffa du caoutchouc en l'absence d'air et obtint un liquide appelé *isoprène.* Sa molécule est composée de cinq atomes de carbone et de huit atomes d'hydrogène, disposés de la manière suivante :



Un autre type de sève (*le latex*), obtenu à partir de certains arbres d'Asie du Sud-Est, donne une substance appelée *gutta-percha.* Celle-ci n'a pas l'élasticité du caoutchouc, mais quand elle est chauffée en l'absence d'air, elle produit aussi de l'isoprène.

Le caoutchouc et la gutta-percha sont tous les deux faits de millions d'unités d'isoprène. Comme dans le cas de l'amidon et de la cellulose, la différence entre eux repose sur le mode de liaison. Dans le caoutchouc, les unités d'isoprène sont liées sur le mode « ....uuuuuuu.... » et de telle façon qu'elles forment un serpentin, qui peut s'étirer, permettant un allongement. Dans le cas de la gutta-percha, les unités sont liées de la manière « ....unununu.... », et forment des chaînes qui sont plus droites dès le départ et donc beaucoup moins extensibles (figure 11.3).

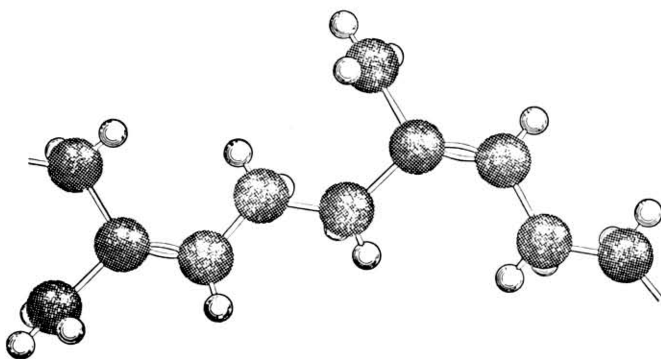


Figure 11.3. La molécule de gutta-percha, dont seulement une partie est montrée ici, est constituée de milliers d'unités d'isoprène. Les cinq premiers atomes de carbone *à gauche* (les boules noires), ainsi que les huit atomes d'hydrogène qui leur sont liés, constituent une unité d'isoprène.

Une seule molécule de sucre (ou ose), telle que le glucose, est un *monosaccharide* (du grec pour « un sucre »). Le saccharose et le lactose sont des *disaccharides* (« deux sucres ») et l'amidon et la cellulose sont des *polysaccharides* (« nombreux sucres »). De même, deux molécules d'isoprène s'unissent pour donner un type de composé bien connu appelé le *terpène* (obtenu à partir de la térébenthine) ; le caoutchouc et la gutta-percha sont appelés polyterpènes.

Le terme général pour ces composés a été inventé, dès 1830, par Berzelius (un grand inventeur de noms et de symboles). Il a appelé l'unité de base un *monomère* (« une partie ») et la grande molécule, un *polymère* (« nombreuses parties »). Les polymères constitués de nombreuses unités (plus de cent par exemple) sont maintenant appelés *hauts polymères*. L'amidon, la cellulose, le caoutchouc et la gutta-percha sont tous des exemples de hauts polymères.

Les polymères ne sont pas des composés bien définis, mais plutôt des mélanges complexes de molécules de différentes tailles. Leur poids moléculaire moyen peut être déterminé par différentes méthodes, dont l'une est la mesure de la *viscosité* (la facilité ou la difficulté avec laquelle un liquide s'écoule sous une pression donnée). Plus les molécules sont grandes et plus elles sont allongées, plus leur participation au *coefficient de friction interne* du liquide est grande, et plus le liquide s'écoule comme un sirop, et non comme de l'eau. Le chimiste allemand Hermann Staudinger développa cette méthode en 1930, dans le cadre de son travail général sur les polymères, et reçut le prix Nobel de chimie pour sa contribution à la compréhension de ces molécules géantes.

En 1913, deux chimistes japonais découvrirent que les fibres naturelles telles que celles de la cellulose diffractent les rayons X, comme le font les cristaux. Les fibres ne sont pas des cristaux dans le sens habituel du terme, mais elles ont le caractère de *microcristaux*. Les longues chaînes d'unités constituant la cellulose ont tendance par endroits à courir en faisceaux parallèles sur des distances plus ou moins longues. Tout au long

de ces faisceaux parallèles, les atomes sont arrangés en ordre répétitif comme ils le seraient dans des cristaux, et les rayons X frappant ces sections de la fibre sont diffractés.

C'est ainsi qu'on en est venu à séparer les polymères en deux grandes catégories : *cristalline et amorphe.*

Dans un polymère cristallin, tel que la cellulose, la cohésion des chaînes individuelles est augmentée par le fait que les chaînes voisines et parallèles sont liées entre elles par des liaisons chimiques. Les fibres résultantes ont une résistance de tension considérable. L'amidon est cristallin aussi, mais à un degré nettement inférieur à celui de la cellulose, et n'a donc pas la résistance de cette dernière, ni sa capacité à former des fibres.

Le caoutchouc est un polymère amorphe, car il n'y a pas de formation de liaisons inter-chaînes. Quand on chauffe, les différentes chaînes peuvent bouger indépendamment et glisser librement les unes sur les autres, ou les unes autour des autres. En conséquence, le caoutchouc ou les polymères caoutchouteux apparentés deviennent mous et collants et finissent par fondre à la chaleur. En allongeant le caoutchouc, on redresse les chaînes et on introduit ainsi un certain caractère microcristallin. Le caoutchouc étiré a donc une force de tension considérable. La cellulose et l'amidon, dans lesquels les molécules individuelles sont liées entre elles ici et là, ne peuvent pas avoir la même liberté de mouvement, et il n'y a donc pas d'amollissement à la chaleur. Ils restent rigides jusqu'à ce que la température soit suffisamment élevée pour induire des vibrations qui ébranlent la molécule jusqu'à la casser, causant la carbonisation et le dégagement de fumée.

A des températures inférieures au stade visqueux et collant, les polymères amorphes sont souvent mous et élastiques. Cependant, à des températures encore plus basses, ils deviennent durs et raides, et même cassants comme du verre. Le caoutchouc naturel n'est sec et élastique que dans une zone de températures assez limitée. L'addition de soufre à concurrence de 5 à 8 % crée des liaisons sulfure souples entre les chaînes, qui réduisent l'indépendance des chaînes et évitent donc la viscosité à des températures modérées. Elles accroissent aussi le jeu libre entre les chaînes à des températures moyennement basses, empêchant ainsi le caoutchouc de durcir. L'addition de soufre en plus grandes quantités, jusqu'à 30 ou 50 %, lie les chaînes si fortement que le caoutchouc devient dur ; il porte alors le nom de *caoutchouc dur ou ébonite.*

Même le caoutchouc vulcanisé devient comme du verre si la température est suffisamment réduite. Une balle de caoutchouc ordinaire, trempée quelques instants dans l'air liquide, se brise en éclats quand on la lance contre un mur. C'est une des démonstrations préférées des cours d'introduction à la chimie.

Des polymères amorphes différents ont des propriétés physiques différentes à une température donnée. A la température ambiante, le caoutchouc naturel est élastique, d'autres résines sont vitreuses et solides, et le chiclé (obtenu à partir du sapotillier d'Amérique du Sud) est mou et visqueux (c'est l'ingrédient principal du chewing-gum).

LA CELLULOSE ET LES EXPLOSIFS

A part la nourriture, principalement constituée de hauts polymères, le polymère dont l'homme a dépendu le plus longtemps est probablement

la cellulose. C'est le principal composant du bois, qui a été depuis toujours un combustible et un matériau de construction indispensables. La cellulose est aussi utilisée pour faire du papier. Constituant la fibre du coton et du lin, la cellulose est le produit textile le plus important. Aussi les chimistes du milieu du XIXe siècle se sont-ils tournés vers la cellulose comme source de matière première pour fabriquer d'autres molécules géantes.

On peut modifier la cellulose en fixant à la combinaison oxygène-hydrogène (groupements *hydroxyles*) des unités de glucose, un *groupement nitrate* (un atome d'azote et trois atomes d'oxygène). Cela se fait en traitant la cellulose avec un mélange d'acide nitrique et sulfurique ; on a créé ainsi un explosif d'une puissance encore inconnue. Cet explosif a été découvert par hasard en 1846 par le chimiste allemand, né en Suisse, Christian Schönbein (qui, en 1839, avait découvert l'ozone). Après avoir renversé un mélange d'acide dans sa cuisine, il prit le tablier en coton de sa femme, ainsi que le raconte l'histoire, pour éponger le désastre. Quand il le suspendit pour le faire sécher au-dessus du feu, le tablier explosa sans laisser de trace.

Schönbein envisagea immédiatement les potentialités de sa découverte, comme l'indique le nom qu'il donna au composé qui, traduit en français, serait *coton-poudre*. On l'appelle aussi *nitrocellulose*. Schönbein proposa la recette à plusieurs gouvernements. La poudre à canon ordinaire émettait tant de fumée que les canonniers étaient noircis, les canons encrassés devaient être écouvillonnés entre les tirs, et le voile de fumée était si épais qu'après les premières salves, on finissait par se battre à l'aveuglette. Les ministères de la guerre sautèrent donc sur l'occasion qui leur était offerte d'utiliser un explosif non seulement puissant, mais qui ne faisait pas de fumée. Le coton-poudre était un explosif impatient, qui n'attendait pas d'être dans le canon pour exploser. Au début des années 1860, le boom du coton-poudre était terminé, au sens propre comme au sens figuré.

Néanmoins, plus tard, on découvrit des méthodes permettant d'éliminer les impuretés présentes en petites quantités dans le coton-poudre, qui provoquaient son explosion. Il devint alors d'un maniement plus facile. Le chimiste anglais Dewar (rendu célèbre par la liquéfaction des gaz) et un de ses collaborateurs, Frederick Augustus Abel, mélangèrent en 1889 de la *nitroglycérine* et du coton-poudre, et ajoutèrent de la vaseline au mélange pour permettre son moulage sous forme de cordon (le mélange était appelé *cordite*). Enfin, on obtenait une poudre, sans fumée et utilisable. La guerre hispano-américaine de 1898 fut la dernière guerre importante où l'on s'est battu avec de la poudre à fusil ordinaire.

L'ère de la machine apporta sa contribution aux horreurs de l'artillerie. Dans les années 1860, l'inventeur américain Richard Gatling fabriqua le premier *fusil mitrailleur* qui permettait de tirer rapidement un grand nombre de balles ; et il fut perfectionné par un autre inventeur américain, Hiram Stevens Maxim, dans les années 1880. La mitrailleuse (*fusil de Gatling*) donna naissance au mot argot *gat* pour désigner un fusil dans l'armée américaine. Ce dernier, ainsi que le *fusil de Maxim*, donnèrent aux impérialistes cyniques de la fin du XIXe siècle un avantage sans précédent sur les « races inférieures » d'Afrique et d'Asie, pour utiliser la phrase blessante de Rudyard Kipling. Un refrain populaire disait : « Quoi qu'il advienne nous avons / Le fusil de Maxim et ils ne l'ont pas ! »

On a continué à progresser ainsi au cours du XXe siècle. L'explosif le plus utilisé durant la Première Guerre mondiale était le *trinitrotoluène,*

généralement abrégé en TNT. Pendant la Deuxième Guerre mondiale, un explosif encore plus puissant, la *cyclonite*, fut utilisé. Tous ces explosifs contiennent le groupement nitro (NO_2) au lieu du groupement nitrate (ONO_2). Ils ouvrirent la voie à la bombe atomique de 1945 (voir chapitre 10).

La nitroglycérine fut découverte au même moment que le coton-poudre. Un chimiste italien, Asciano Sobrero, traita du glycérol avec un mélange d'acide nitrique et sulfurique et comprit qu'il avait découvert quelque chose en risquant de se tuer dans l'explosion qui s'ensuivit. Sobrero, n'ayant pas les talents de promoteur de Schönbein, pensa que la nitroglycérine était une substance trop dangereuse pour être utilisée, et censura pratiquement toute information à son sujet. Mais moins de dix ans plus tard, une famille suédoise, les Nobel, décida d'en fabriquer pour servir « d'huile de mine » dans les travaux de construction et dans les mines. Après une série d'accidents, dont un qui tua un membre de la famille, Alfred Bernhard Nobel, le frère de la victime, mit au point une technique nouvelle : il mélangea la nitroglycérine à une poudre absorbante appelée *kieselguhr* ou *terre de diatomées* (la kieselguhr contient principalement de tout petits squelettes d'organismes unicellulaires, les diatomées). Le mélange se composait de trois parts de nitroglycérine pour une de kieselguhr, mais le pouvoir absorbant de cette dernière était tel que le mélange devenait essentiellement une poudre sèche. On pouvait laisser tomber, frapper avec un marteau et même brûler un bâton de cette terre imprégnée (*dynamite*) sans qu'il explose, mais quand on amorçait la dynamite électriquement et à distance, elle déployait toute la force de la nitroglycérine pure.

Les amorces contiennent des explosifs qui éclatent sous l'effet de la chaleur ou d'un choc mécanique, et elles sont donc appelées *détonateurs*. Le choc puissant de la détonation fait exploser la dynamite. Il pourrait sembler que le danger ait été transféré de la nitroglycérine au détonateur, mais ce n'est pas tellement dramatique puisqu'on n'utilise les détonateurs qu'en toutes petites quantités. Les détonateurs les plus utilisés sont le fulminate de mercure ($HgC_2N_2O_2$) et l'azide de plomb (PbN_6).

Les bâtons de dynamite ont par la suite permis de creuser dans l'Ouest américain des chemins de fer, des mines, des routes et des barrages à un rythme sans précédent dans l'histoire. La dynamite, ainsi que les autres explosifs qu'il découvrit, firent de Nobel, personnage solitaire et impopulaire, un multimillionnaire (qui, malgré sa volonté humanitaire, fut considéré comme un « marchand de mort »). Quand il mourut en 1896, il laissa derrière lui un fonds qui permit de distribuer chaque année les célèbres prix Nobel dans cinq disciplines : chimie, physique, médecine et physiologie, littérature et paix ; et depuis 1969, sciences économiques. A la distinction purement honorifique, s'ajoute une récompense en argent. Les premiers prix ont été décernés en 1901, année du cinquième anniversaire de la mort de Nobel, et ils sont devenus le plus grand honneur auquel un chercheur puisse prétendre.

Considérant la nature de la société humaine, de grands savants continuent à consacrer une fraction non négligeable de leur temps aux explosifs. Presque tous les explosifs contiennent de l'azote ; la chimie de cet élément et de ses composés est donc d'une importance clef (elle est aussi d'une importance clef pour le monde vivant, il faut bien l'admettre).

Le chimiste allemand Willhem Ostwald, qui s'intéressait davantage à la théorie de la chimie qu'aux explosifs, étudia la vitesse à laquelle les

réactions chimiques s'effectuent. Il appliqua à la chimie les principes mathématiques utilisés en physique, et devint ainsi l'un des fondateurs de la *chimie physique.* Vers la fin du siècle il élabora de nouvelles méthodes pour convertir l'ammoniac (NH_3) en oxydes d'azote, qui pouvaient alors être utilisés pour fabriquer des explosifs. En récompense de son travail théorique, en particulier sur la catalyse, Ostwald reçut le prix Nobel de chimie en 1909.

Au début du XXe siècle, une des dernières sources d'azote utilisables était les dépôts de nitrates découverts dans le désert, au nord du Chili. Pendant la Première Guerre mondiale, ces mines furent rendues inaccessibles aux Allemands par la marine anglaise. Pour pallier ce manque, le chimiste allemand Fritz Haber mit au point une technique par laquelle l'azote moléculaire de l'air peut être combinée à l'hydrogène sous pression, pour former l'ammoniac nécessaire au procédé de Ostwald. Le *procédé de Haber* fut amélioré par le chimiste allemand Karl Bosch, qui supervisa la construction d'usines pour fabriquer l'ammoniac pendant la guerre. Haber reçut le prix Nobel de chimie en 1918 et Bosch le partagea en 1931. A la fin des années soixante, les États-Unis fabriquaient à eux seuls douze millions de tonnes d'ammoniac chaque année selon le procédé de Haber.

PLASTIQUES ET CELLULOÏD

Retournons à la cellulose modifiée. De toute évidence, c'était l'addition de groupements nitrates qui créait la nature explosive de la cellulose. Dans le coton-poudre tous les groupements hydroxyles disponibles étaient nitrés. Que se passerait-il si seuls quelques-uns d'entre eux étaient nitrés ? La cellulose modifiée serait-elle moins explosive ? En fait, une cellulose partiellement nitrée s'avéra ne pas être explosive du tout. En échange, elle brûlait très facilement ; ce matériau fut finalement appelé *pyroxyline* (du grec signifiant « feu de bois »).

La pyroxyline peut être dissoute dans un mélange d'alcool et d'éther – comme le découvrirent, chacun de leur côté, un savant français, Louis Nicolas Ménard, et un étudiant en médecine américain, J. Parkers Maynard (n'est-ce pas une curieuse similitude de nom ?). Après évaporation de l'éther et de l'alcool, la pyroxyline se transformait en un film dur et transparent qu'on nomma *collodion.* D'abord utilisé comme pansement sur les petites coupures et écorchures, il fut surnommé *la nouvelle peau.* Cependant, les aventures de la pyroxyline ne faisaient que commencer ; bien d'autres se profilaient à l'horizon.

La pyroxyline brute est, par elle-même, cassante. Mais le chimiste anglais Alexander Parkes découvrit que si on ajoutait une substance telle que le camphre à de la pyroxyline dissoute dans un mélange d'alcool et d'éther, après évaporation du solvant, on obtenait un solide dur qui devenait souple et malléable après chauffage. Il pouvait alors être moulé dans la forme désirée qu'il gardait en durcissant après refroidissement. C'est donc la nitrocellulose qui a permis de fabriquer le premier *plastique* artificiel, en 1865. Le camphre, en conférant des propriétés plastiques à un matériau qui, autrement, est cassant, devenait le premier *plastifiant.*

Les plastiques cessèrent d'être une curiosité chimique et devinrent populaires lors de leur introduction dans les salles de billard. Les boules de billard étaient à l'époque faites d'ivoire, une denrée que l'on ne pouvait

obtenir que des défenses d'éléphants, ce qui, naturellement, posait problème. Au début des années 1860, un prix de dix mille dollars fut offert pour le meilleur succédané de l'ivoire qui remplirait les exigences multiples d'une boule de billard : dureté, élasticité, résistance à la chaleur et à l'humidité, absence de grain, etc. L'inventeur américain John Wesley Hyatt fut de ceux qui concoururent pour le prix. Il commença à progresser lorsqu'il entendit parler de l'astuce de Parkes, qui lui permit de plastifier la pyroxyline en une matière moulable se figeant en un solide dur. Hyatt chercha à améliorer les techniques de fabrication de ce matériau, en utilisant moins d'alcool et d'éther, qui sont coûteux, et plus de chaleur et de pression. Dès 1869, Hyatt produisait à partir de ce matériau, qu'il appela *celluloïd,* des boules de billard à bon marché. Cela lui permit de gagner le prix.

On comprit très vite que le celluloïd avait bien d'autres applications. Il faisait preuve en fait d'une grande souplesse d'emploi. On pouvait le mouler à la température de l'eau bouillante, le couper, le percer et le scier à des températures plus basses. Brut, il était solide et dur, mais pouvait aussi être produit sous la forme d'un film flexible utilisable pour les cols de chemises, les hochets d'enfants, etc. Sous forme de film encore plus mince et plus flexible, on pouvait s'en servir comme support pour des composés d'argent noyés dans la gélatine, et on obtint ainsi le premier film photographique facile d'emploi.

Le grand défaut du celluloïd est que, en raison de ses groupements nitrates, il a tendance à brûler avec une rapidité extrême, en particulier quand il est sous la forme d'un film mince. Il a été à l'origine d'un grand nombre d'incendies tragiques.

Le remplacement des groupements nitrates par des groupements acétates (CH_3COO-) aboutit à la formation d'un autre type de cellulose : l'*acétate de cellulose.* Plastifié correctement, il a des propriétés aussi bonnes ou presque aussi bonnes que celles du celluloïd avec, en plus, l'énorme avantage d'être bien moins inflammable. L'acétate de cellulose a commencé à être utilisé juste avant la Première Guerre mondiale ; et après la guerre, il a remplacé complètement le celluloïd dans la fabrication des films photographiques et de nombreux autres articles.

LES HAUTS POLYMÈRES

Au cours du demi-siècle qui suivit le développement du celluloïd, les chimistes s'émancipèrent de la dépendance vis-à-vis de la cellulose comme base de fabrication des plastiques. Dès 1872, Baeyer (qui devait synthétiser plus tard l'indigo) avait remarqué qu'en chauffant ensemble phénols et aldéhydes, on obtenait une masse résineuse et gluante. Comme il n'était intéressé que par les petites molécules qu'il pouvait isoler du milieu réactionnel, il négligea ce bourbier au fond du ballon (les chimistes organiciens du XIXe siècle avaient eux aussi tendance à le faire quand une matière gluante encrassait leur verrerie). Trente-sept ans plus tard, le chimiste américain d'origine belge Leo Hendrick Baekeland, au cours d'expériences avec le formaldéhyde, fit la découverte suivante : dans certaines conditions, la réaction produisait une résine qui, si on continuait à la chauffer sous pression, devenait d'abord un solide mou, puis une substance dure et insoluble. Cette résine, on pouvait soit la mouler, pendant

qu'elle était molle, et elle se figeait ensuite en une forme permanente et solide ; soit, une fois dure, la réduire en poudre, la verser dans un moule et en faire une pièce en chauffant sous pression. Des formes très complexes peuvent être moulées facilement et rapidement. De plus, le produit est inerte et résiste à la plupart des agressions de l'environnement.

Baekeland appela son produit Bakélite, d'après son propre nom. La bakélite appartient à la classe des plastiques *thermodurcissables,* qui une fois figés par refroidissement, ne peuvent plus être ramollis par chauffage (néanmoins, ils peuvent être détruits par une chaleur intense). Des matériaux tels que les dérivés de la cellulose, que l'on peut ramollir maintes et maintes fois, sont appelés *thermoplastiques.* La bakélite a de nombreuses utilisations – en tant qu'isolant, adhésif, agent de lamination, etc. Bien qu'elle soit le plus ancien des plastiques thermodurcissables, c'est toujours le plus utilisé.

La bakélite a été le premier haut polymère produit en laboratoire à partir de petites molécules. Pour la première fois, le chimiste a dirigé cette opération d'un bout à l'autre. Cela ne représente pas, naturellement, une synthèse telle qu'on l'entend pour la quinine ou l'hème, où le chimiste doit placer chaque nouvel atome dans sa position exacte, presque un par un. Au contraire, la production de hauts polymères requiert simplement que les petites unités dont ils sont composés soient mélangées dans des conditions convenables. Une réaction est alors amorcée, dans laquelle les unités forment spontanément une chaîne sans l'intervention ponctuelle du chimiste. Le chimiste peut modifier la nature de la chaîne indirectement en changeant les produits de départ et leurs proportions, ou en ajoutant de faibles quantités d'acides, de bases ou de substances variées, qui agissent comme *catalyseurs* en orientant avec précision la nature de la réaction.

Après le succès de la bakélite, les chimistes se tournèrent naturellement vers d'autres matières premières, à la recherche de nouveaux hauts polymères synthétiques qui pourraient se révéler utiles comme plastiques. Le succès leur sourit plus d'une fois.

Les chimistes britanniques découvrirent, dans les années trente, que le gaz éthylène ($CH_2 = CH_2$), à chaud et sous pression, formait de très longues chaînes. L'une des deux liaisons de la double liaison entre les atomes de carbone s'ouvre et s'attache à la molécule voisine. Cela se passant de proche en proche, le résultat est une molécule constituée d'une longue chaîne appelée *polyéthylène.*

La molécule de la cire de paraffine est une longue chaîne faite des mêmes unités, mais la molécule de polyéthylène est encore plus longue. Le polyéthylène est donc une supercire. Il a de la cire la blancheur laiteuse, la sensation glissante, les propriétés d'isolant électrique, l'imperméabilité à l'eau et la légèreté (c'est à peu près le seul plastique qui puisse flotter sur l'eau). De plus, quand il est de la meilleure qualité, il est bien plus résistant et flexible que la paraffine.

Au début de la fabrication du polyéthylène, il fallait utiliser des pressions élevées, donc dangereuses, et le produit avait un point de fusion assez bas – juste au-dessous du point d'ébullition de l'eau. Il mollissait au point d'être inutilisable à des températures inférieures à son point de fusion. Cette propriété était apparemment due au fait que les chaînes de carbone étaient ramifiées, empêchant les molécules de s'assembler en des faisceaux serrés et cristallins. En 1953, un chimiste allemand nommé Karl Ziegler trouva un moyen de produire des chaînes de polyéthylène non ramifiées,

sans que de hautes pressions soient nécessaires. On obtenait ainsi une nouvelle variété de polyéthylène, plus résistante et plus forte que la première, et capable de supporter la température de l'eau bouillante sans trop se ramollir. Ziegler réalisa cela en utilisant un nouveau type de catalyseur : une résine contenant des ions métalliques, tels que l'aluminium ou le titane, fixés le long de la chaîne à des groupements chargés négativement.

Le chimiste italien Giulio Natta, ayant entendu parler des catalyseurs organométalliques de Ziegler, appliqua la technique à la condensation du *propylène* (de l'éthylène auquel un groupement méthyle de un carbone, CH_3-, est attaché, soit $CH_2 = CH - CH_3$). En moins de dix semaines, il avait trouvé que, dans le polymère résultant, tous les groupements méthyles étaient orientés dans la même direction, au lieu d'être orientés d'une façon aléatoire dans toutes les directions, comme c'était le cas auparavant dans la formation des polymères. De tels polymères isotactiques (un nom proposé par la femme de Natta) se révélèrent présenter des propriétés intéressantes, et on peut maintenant les fabriquer à volonté. Les chimistes peuvent définir les polymères avec une précision jamais atteinte. Pour leurs travaux sur ce sujet, Ziegler et Natta partagèrent le prix Nobel de chimie en 1963.

Au cours de la mise au point de la bombe atomique, on découvrit un autre haut polymère utile, sous la forme d'un curieux parent du polyéthylène. Pour séparer l'uranium 235 de l'uranium naturel, les physiciens nucléaires doivent faire réagir l'uranium avec le fluor pour former l'hexafluorure d'uranium. Le fluor est un des éléments les plus réactifs et s'attaque à peu près à tout. Cherchant pour leurs récipients des lubrifiants et des joints qui résisteraient à l'attaque du fluor, les physiciens eurent recours aux *fluorocarbones* – des susbstances dans lesquelles quelques hydrogènes sont remplacés par des fluors.

Jusque-là, les fluorocarbones n'avaient été que des curiosités de laboratoire. La première (et la plus simple) de cette classe de molécules, le *tétrafluorure de carbone* (CF_4), n'avait été obtenue sous forme pure qu'en 1926. A cette époque, on étudiait la chimie de ces substances avec intensité. Parmi les fluorocarbones étudiés, il y avait le *tétrafluoroéthylène* ($CF_2 = CF_2$), synthétisé pour la première fois en 1933 ; c'est, comme vous le voyez, de l'éthylène où les quatre hydrogènes ont été remplacés par quatre fluors. Fatalement, un jour, on aurait l'idée de polymériser le tétrafluoroéthylène au même titre que l'éthylène. Après la guerre, les chimistes de Du Pont de Nemours produisirent un polymère à longue chaîne qui était uniformément constitué de ....$CF_2CF_2CF_2CF_2$...., très voisin du polyéthylène qui est ....$CH_2CH_2CH_2CH_2CH_2$.... Son nom de marque est le Téflon (*tefl* étant une abréviation de *tétrafluoro*).

Le téflon ressemble au polyéthylène, en mieux. Les liaisons carbone-fluor sont plus fortes que les liaisons carbone-hydrogène et moins sensibles à l'agression de l'environnement. Le téflon n'est soluble dans rien, et rien ne peut le mouiller ; c'est un excellent isolant électrique et il est considérablement plus résistant à la chaleur que même le nouveau polyéthylène amélioré. L'application la plus connue est le revêtement des poêles à frire qui permet de faire cuire les aliments sans graisse, puisque la nourriture ne peut pas coller à ce polymère de fluorocarbone peu liant.

Un composé intéressant, qui n'est pas tout à fait un fluorocarbone, est le fréon (CF_2CL_2), déjà mentionné. Il a été introduit en 1932 en tant que

refrigérant. Il est plus cher que l'ammoniac ou l'anhydride sulfureux, mais il est utilisé dans les congélateurs à l'échelle industrielle, car il est aussi inodore, non toxique, et ininflammable. Si une fuite accidentelle se produit, elle ne provoque qu'un danger minime. Midgley, qui le mit au point, démontra son innocuité en aspirant une grande bouffée de fréon qu'il expira lentement au-dessus de la flamme d'une bougie. La bougie s'éteignit, mais Midgley était indemne. C'est grâce au fréon que l'air conditionné est devenu partie intégrante de la vie américaine depuis la Deuxième Guerre mondiale.

LE VERRE ET LE SILICONE

Les propriétés plastiques n'appartiennent pas, bien sûr, qu'au seul monde organique. L'une des substances plastiques les plus anciennes est le verre. Les grandes molécules de verre sont essentiellement faites de chaînes d'atomes de silicium et d'oxygène : Si-O-Si-O-Si-O-Si-O, et ainsi de suite, indéfiniment. Les atomes de silicium, comme ceux de carbone, ont quatre liaisons de valence, donc chaque atome de silicium dans la chaîne a deux liaisons libres auxquelles on peut attacher d'autres groupements. La liaison silicium-silicium est plus faible que la liaison carbone-carbone, de sorte qu'on ne peut former que de courtes chaînes de silicium, et celles-ci (appelées *silanes*) sont instables. En revanche, la liaison silicium-oxygène est forte et de telles chaînes sont encore plus stables que celles de carbone. En fait, étant donné que la croûte terrestre est par moitié constituée d'oxygène et pour un quart de silicium, la terre ferme sur laquelle nous nous tenons peut être considérée comme étant essentiellement une chaîne silicium-oxygène.

Bien que les beautés et les utilisations du verre (une sorte de sable rendu transparent) soient infinies, il présente le grand inconvénient d'être cassable, et lorsqu'on le brise il donne des morceaux durs et coupants qui sont dangereux et même mortels. Si le pare-brise d'une voiture est fait de verre non traité, une collision peut transformer une voiture en bombe shrapnel.

On peut cependant préparer deux feuilles de verre entre lesquelles est placée une fine couche de polymère transparent qui durcit en agissant comme adhésif. C'est le *verre securit :* lorsqu'il est brisé en éclats, même de taille infime, chaque morceau est maintenu fermement en place par le polymère, aucun d'eux ne vole. A l'origine, dès 1905, le collodion était utilisé comme liant, mais il est remplacé de nos jours, la plupart du temps, par des polymères fabriqués à partir de petites molécules telles que le chlorure de vinyle (le chlorure de vinyle est de l'éthylène où un des atomes d'hydrogène est remplacé par un atome de chlore). La *résine de vinyle* est incolore et le reste ; on peut donc être sûr que le verre securit ne jaunira pas avec le temps.

Enfin, il y a des plastiques transparents qui peuvent complètement remplacer le verre, du moins dans certaines applications. Au milieu des années trente, Du Pont polymérisa une petite molécule appelée le *méthacrylate de méthyle,* et forma le polymère qui en résultait (un plastique *polyacrylique*) en plaques transparentes et claires. Les marques de ces produits sont Plexiglas et Lucite. Ces *verres organiques* sont plus légers que le verre ordinaire, plus facilement moulés, moins fragiles, et se cassent

nettement au lieu de voler en éclats. Pendant la Deuxième Guerre mondiale, on utilisa des plaques de plastique transparentes et moulées comme fenêtres dans les avions, où la légèreté et la solidité sont particulièrement appréciées. Bien sûr, les plastiques polyacryliques ont leurs inconvénients. Ils sont sensibles aux solvants organiques, ils sont plus facilement amollis par la chaleur que le verre, et ils se rayent facilement. Si les plastiques acryliques étaient utilisés comme pare-brise sur les voitures, par exemple, ils se rayeraient rapidement sous l'effet des particules de poussière et deviendraient dangereusement dépolis. En conséquence, le verre n'est pas à la veille de se voir complètement remplacé. En fait, on lui trouve même de nouvelles applications. Des fibres de verre ont été filées en un matériau textile qui a toute la souplesse des fibres organiques et offre l'avantage inestimable de résister totalement au feu.

En plus des succédanés du verre, il y a ce que l'on pourrait appeler les compromis du verre. Comme je l'ai dit, dans une chaîne silicium-oxygène, chaque atome de silicium a deux liaisons qui peuvent se lier à d'autres atomes. Dans le verre, ces autres atomes sont des atomes d'oxygène, mais il n'est pas nécessaire qu'il en soit ainsi. Que se passerait-il si des groupements contenant des atomes de carbone étaient attachés à la place de l'oxygène ? On aurait alors, en quelque sorte, une chaîne minérale avec des ramifications organiques – un compromis entre des matériaux organique et minéral. Dès 1908, le chimiste anglais Frederic Stanley Kipping avait formulé de tels composés, connus sous le nom de *silicones*.

Pendant la Deuxième Guerre mondiale, les longues chaînes de *résine de silicone* devinrent un matériau de choix. Ces silicones sont plus résistants à la chaleur que les polymères purement organiques. En variant la longueur de la chaîne et la nature des chaînes latérales, on peut obtenir une liste de propriétés souhaitées que le verre ne possède pas. Par exemple, certains silicones sont liquides à la température ambiante et ne changent que très peu de viscosité sur de grands écarts de température, c'est-à-dire qu'ils ne se fluidifient pas à la chaleur et ne s'épaississent pas au froid. C'est une propriété particulièrement utile pour les fluides hydrauliques – comme ceux utilisés pour descendre le train d'atterrissage des avions, par exemple. D'autres silicones forment des joints souples ressemblant au mastic, qui ne durcissent pas, ne se fendillent pas et sont remarquablement imperméables à l'eau. D'autres encore servent de lubrifiants résistant à l'acide, etc.

Les fibres synthétiques

Les fibres synthétiques constituent un chapitre particulièrement intéressant de l'histoire de la synthèse organique. La cellulose servit de matière première dans la fabrication des premières fibres synthétiques (comme des premiers plastiques à l'état brut). Naturellement, les chimistes utilisèrent d'abord le nitrate de cellulose, qui était disponible en quantités suffisantes. En 1844, Hilaire Bernigaud de Chardonnet, un chimiste français, a dissous du nitrate de cellulose dans un mélange d'alcool et d'éther, et fait passer

la solution épaisse résultante à travers de petits trous. Tandis que la solution sortait par les trous, l'alcool et l'éther s'évaporaient, laissant derrière eux un fil de collodion très fin. (C'est ainsi que les araignées et les vers à soie fabriquent leur fil : ils éjectent, à travers des orifices minuscules, un liquide qui devient un fil solide à la suite de son contact avec l'air.) Les fibres de nitrate de cellulose étaient trop inflammables pour être utilisées, mais on pouvait éliminer les groupements nitrates par un traitement chimique approprié. Le résultat fut un fil de cellulose brillant, qui ressemblait à de la soie.

Le procédé de Chardonnet était évidemment onéreux, puisqu'on commençait par fixer des groupements nitrates pour les enlever ensuite. Sans parler du dangereux intervalle pendant lequel ces groupements sont encore en place et du mélange éther-alcool utilisé comme solvant, qui est lui aussi dangereusement inflammable. On a réussi à solubiliser la cellulose en 1892. Le chimiste anglais Charles Frederick Cross en a dissous, par exemple, dans du carbone sulfure, et a fabriqué une fibre à partir de la solution visqueuse ainsi obtenue (la *viscose*). Le problème est que le carbone sulfure est inflammable et toxique. En 1903, on a commencé à utiliser un procédé concurrent utilisant un solvant dont l'acide acétique était l'un des constituants, et à former l'*acétate de cellulose.*

Ces fibres artificielles furent appelées d'abord *soie artificielle,* puis *rayonne,* à cause de leur aspect brillant qui reflète les rayons de la lumière. On distingue deux variétés principales de rayonne : la *rayonne viscose* et la *rayonne acétate.*

On peut faire passer la viscose à travers une fente et former ainsi un film flexible, imperméable et transparent – la *cellophane.* Ce procédé a été inventé en 1908 par le chimiste français Jacques Edwin Brandenberger. On peut agir de même, dans le même but, avec d'autres polymères synthétiques. Les résines vinyliques, par exemple, produisent le film de recouvrement connu sous le nom de Saran.

La première fibre totalement synthétique a vu le jour dans les années trente.

Parlons d'abord de la soie. La soie est un produit d'origine animale, faite par des chenilles qui sont très exigeantes en ce qui concerne leurs soins et leur nourriture. La fibre doit être minutieusement déroulée des cocons. C'est pour cela que la soie est chère et ne peut être produite à grande échelle. Elle l'a d'abord été, il y a plus de deux mille ans en Chine, où le secret de sa production était jalousement gardé, afin de conserver un monopole lucratif à l'exportation. Cependant, les secrets ne peuvent pas être éternellement gardés, malgré toutes les mesures de sécurité. Le secret se répandit en Corée, au Japon et aux Indes. La soie était acheminée vers la Rome antique par la longue route terrestre, à travers l'Asie. Les intermédiaires levaient un octroi à chaque étape de la route ; la fibre n'était donc accessible qu'aux plus riches. En 550 av. J.-C., des œufs de vers à soie furent apportés en fraude à Constantinople, et la production de soie démarra en Europe. La soie est néanmoins restée plus ou moins un objet de luxe. De plus, récemment encore, il n'y avait pas de bon produit de remplacement. La rayonne en imite le chatoiement, mais ni la finesse ni la résistance.

Après la Première Guerre mondiale, quand les bas de soie devinrent un accessoire indispensable de la garde-robe des femmes, le besoin de disposer de davantage de soie ou d'un succédané adéquat se fit fortement sentir. C'était particulièrement vrai aux États-Unis, où la soie était utilisée en

quantités plus grandes qu'ailleurs, et où les relations avec le fournisseur principal, le Japon, se détérioraient progressivement. Les chimistes rêvèrent d'un moyen de fabriquer une fibre qui pourrait se comparer à la soie.

La soie est une protéine (voir chapitre 12). Sa molécule est constituée de monomères appelés *acides aminés*, qui contiennent eux-mêmes un groupement *aminé* (-NH_2) et un groupement *carboxyle* (-COOH). Les deux groupements sont reliés entre eux par un atome de carbone. Si on désigne le groupement aminé par *a*, le groupement carboxyle par *c* et le carbone intermédiaire par un tiret, un acide aminé s'écrit alors : a-c. Ces acides aminés se polymérisent tête-bêche, c'est-à-dire que le groupement aminé de l'un se condense avec le groupement carboxyle du voisin. La structure de la soie est donc une molécule qui peut s'écrire : ....a-c.a-c.a-c.a-c.a-c.....

Dans les années trente, un chimiste de Du Pont, Wallace Hume Carothers, a recherché des molécules contenant des groupements aminés et des groupements carboxyles, dans l'espoir de les faire se condenser pour former des molécules contenant de grands cycles (ces molécules sont importantes pour la parfumerie). Au lieu de cela, il a trouvé qu'elles se condensaient pour former des molécules à longue chaîne.

Carothers ne s'endormit pas sur sa découverte et ne perdit guère de temps à continuer la recherche sur la formation de cycles. En fin de compte, il forma des fibres à partir de l'acide adipique et de l'hexaméthylène diamine. L'acide adipique contient deux groupements carboxyles séparés par quatre atomes de carbone ; il peut donc être représenté par c----c. L'hexaméthylène diamine est constitué de deux groupements aminés séparés par six atomes de carbone, soit : a------a. Quand Carothers mélangea les deux substances, elles se condensèrent pour former le polymère suivant : a------a.c----c.a------a.c----c.a------a..... On remarque que les points où la condensation a lieu ont la configuration « c.a », qui est celle de la soie.

Les premières fibres ainsi synthétisées n'étaient pas très bonnes, car trop fragiles. Carothers décida que le problème provenait de l'eau formée pendant la condensation. Celle-ci empêche la polymérisation de se poursuivre en amorçant une hydrolyse à contre-réaction. Carothers trouva la solution : il s'arrangea pour que la polymérisation ait lieu à faible pression ; l'eau, ainsi, se vaporisait, et il pouvait l'éliminer en la faisant se condenser sur une surface de verre froide, maintenue près du mélange réactionnel et inclinée de façon que l'eau s'écoule hors de l'enceinte (un *alambic*). La polymérisation peut maintenant se prolonger indéfiniment. De belles longues chaînes droites sont ainsi formées ; et en 1935, Carothers avait enfin jeté les bases d'une fibre de rêve.

Le polymère formé à partir de l'acide adipique et de l'hexaméthylènediamine était fondu puis extrudé à travers des trous. Il était alors étiré de manière à former des fibres qui se juxtaposent en faisceaux cristallins. Le résultat était un fil brillant, ressemblant à la soie, qui pouvait être tissé en un matériau presque aussi beau et diaphane que la soie, même plus solide. Cette première fibre totalement synthétique fut appelée *Nylon*. Carothers est décédé en 1937, avant que sa découverte ne soit complètement exploitée.

Du Pont annonça la naissance de la fibre synthétique en 1938 et commença à la produire commercialement en 1939. Pendant la Deuxième Guerre mondiale, l'armée américaine réquisitionna toute la production de nylon pour fabriquer des parachutes et pour des centaines d'autres usages. Après la guerre, le nylon remplaça complètement la soie dans la fabrication des bas, qui sont appelés communément *bas nylon*.

Le nylon ouvrit la voie à la production de nombreuses autres fibres synthétiques. On peut polymériser l'acrylonitrile, ou cyanure de vinyle ($CH_2 = CHCN$), en une longue chaîne comme celle du polyéthylène, mais avec des groupements nitriles (CN) (parfaitement inoffensifs dans ce cas-là) attachés à un carbone sur deux. Le résultat, qui date de 1950, est l'Orlon. Si le chlorure de vinyle ($CH_2 = CHCl$) est ajouté, de telle façon que la chaîne éventuelle contienne autant d'atomes chlores que de nitriles, on aboutit au Dynel. D'autre part, l'addition de groupements acétates, en utilisant l'acétate de vinyle ($CH_2 = CHOOCH_3$), produit les fibres acryliques (Acrilan).

Les Anglais, en 1941, ont fabriqué une fibre de *polyester,* dont le groupement carboxylique d'un monomère se condense avec le groupement hydroxyle de l'autre. Le résultat est l'habituelle longue chaîne d'atomes de carbone, rompue ici par l'insertion périodique d'un atome d'oxygène dans la chaîne. Les Anglais l'appellent Terylene, mais il est appelé Dacron aux États-Unis et en France.

Ces nouvelles fibres synthétiques sont plus imperméables que la plupart des fibres naturelles. Elles résistent donc plus à l'humidité et on ne peut pas les teindre facilement. Elles ne sont pas sujettes à la destruction par les mites ou autres insectes. Certaines sont infroissables et peuvent être utilisées pour préparer des tissus ne nécessitant pas de repassage.

Le caoutchouc synthétique

Il est difficile d'imaginer que les hommes ne circulent sur des roues en caoutchouc que depuis environ cent ans. Pendant des millions d'années ils ont utilisé des roues en bois ou en fer. Après la découverte de Goodyear, permettant de disposer de caoutchouc vulcanisé, nombreux furent ceux qui pensèrent que l'on pouvait recouvrir les roues de caoutchouc, plutôt que de métal. En 1845, la meilleure idée vint d'un ingénieur anglais, Robert William Thomson : il breveta un dispositif consistant en une chambre à air qui s'adaptait à une roue. En 1890, les pneus étaient couramment employés pour les bicyclettes ; et en 1895, pour les voitures sans chevaux.

Aussi surprenant que cela puisse paraître, le caoutchouc, bien que souple et relativement mou, s'est avéré beaucoup plus résistant à l'abrasion que le bois ou le métal. Sa longévité, liée à ses qualités d'absorption des chocs et associée à la conception de la chambre à air, apportait une amélioration sans précédent au confort du voyage.

Au fur et à mesure que la production automobile augmentait, le besoin de caoutchouc pour les pneus devenait astronomique. En un demi-siècle, la production mondiale de caoutchouc a été multipliée par quarante. On peut juger de la quantité de caoutchouc utilisée pour fabriquer les pneus de nos jours, si l'on sait qu'aux États-Unis, l'ensemble des automobiles usent au moins 200 000 tonnes de caoutchouc sur les routes chaque année, même si, individuellement, chaque voiture en use relativement peu.

La demande accrue de caoutchouc a introduit une certaine insécurité dans l'approvisionnement de nombreuses nations pendant les guerres. Celles-ci s'étant mécanisées, les armées commencèrent à circuler « sur caoutchouc » ; or, le caoutchouc ne pouvait être obtenu en quantité suffisante que de la

Malaisie, très éloignée des nations « civilisées » les plus aptes à s'engager dans des guerres « civilisées ». La Malaisie n'est pas l'habitat naturel de l'arbre à caoutchouc, mais l'arbre y a été transplanté avec succès du Brésil, où le réservoir d'arbres à caoutchouc diminuait régulièrement. L'approvisionnement des États-Unis fut stoppé au début de son entrée dans la Deuxième Guerre mondiale lorsque les Japonais envahirent la Malaisie. C'est pour cette raison que les Américains rationnèrent en premier lieu la fabrication de pneus en caoutchouc, avant même l'attaque de Pearl Harbour.

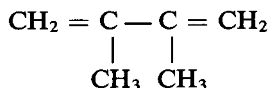
Pendant la Première Guerre mondiale, quand la mécanisation en était à ses débuts, l'Allemagne se trouva gênée, la marine alliée l'ayant coupée de ses sources d'approvisionnement en caoutchouc.

A cette époque, cependant, on envisageait déjà la possibilité de fabriquer du caoutchouc synthétique. Pour produire un tel caoutchouc, il fallait évidemment partir de l'isoprène, l'unité de base du caoutchouc naturel. Dès 1880, les chimistes ont remarqué que l'isoprène, laissé à l'air, avait tendance à devenir collant, et que si on l'acidifiait, il se transformait en une substance ressemblant au caoutchouc. L'empereur Guillaume II, par la suite, utilisa des pneus faits avec ce matériau pour sa voiture officielle, une bonne publicité pour les chimistes allemands.

Quoi qu'il en soit, il y avait deux inconvénients à utiliser l'isoprène comme molécule de départ pour synthétiser le caoutchouc. Le premier est que la source principale d'isoprène est le caoutchouc lui-même. Le deuxième, c'est que lorsque l'isoprène se polymérise, il est plus que vraisemblable qu'il le fait totalement au hasard. Dans le caoutchouc, les chaînes d'isoprène sont toutes enchaînées de la même manière, ...uuuuuuuu.... Dans la gutta-percha, les isoprènes sont enchaînés dans un ordre strictement alterné, ...unununun.... Quand l'isoprène est polymérisé en laboratoire, dans des conditions normales, les *u* et les *n* sont mélangés au hasard, formant un produit qui n'est ni le caoutchouc, ni la gutta-percha. N'ayant ni la souplesse ni la résistance du caoutchouc, il est inutilisable pour la fabrication des pneus.

En fin de compte, des catalyseurs tels que ceux introduits par Ziegler en 1953, pour fabriquer le polyéthylène, permirent de polymériser l'isoprène en un produit presque identique au caoutchouc naturel, mais à cette date, on avait déjà trouvé de nombreux caoutchoucs synthétiques utilisables, ayant des propriétés très différentes de celles du caoutchouc naturel.

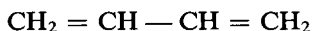
Les premiers efforts se sont naturellement portés sur la fabrication de polymères à partir de composés ressemblant à l'isoprène, et immédiatement disponibles. Par exemple, pendant la Première Guerre mondiale, sous la pression de la pénurie de caoutchouc, l'Allemagne utilisa le diméthylbutadiène :



Contrairement à l'isoprène, qui n'a qu'un seul groupement méthyle (CH_3), le diméthylbutadiène en a deux, un sur chacun des carbones médians de la chaîne de quatre carbones. Le polymère construit à partir du diméthylbutadiène, appelé *méthyl caoutchouc,* pouvait être synthétisé à bon marché et en grandes quantités. L'Allemagne en a produit 2 500 tonnes pendant la Première Guerre mondiale. Il fut le premier caoutchouc synthétique utilisable malgré sa faible résistance à l'effort.

Aux environs de 1930, l'Allemagne et l'Union Soviétique essayèrent une

nouvelle méthode. Elles utilisèrent comme monomère le butadiène, qui n'a aucun groupement méthyle :

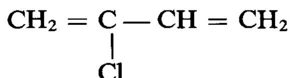


$$CH_2 = CH - CH = CH_2$$

En utilisant du sodium métallique comme catalyseur, elles formèrent un polymère appelé Buna (de *bu*tadiène et *na* pour sodium).

Le caoutchouc buna pouvait être considéré comme satisfaisant en période de restrictions. Il a été amélioré par l'addition d'autres monomères, qui alternent dans la chaîne avec le butadiène. Le meilleur additif a été le styrène, un composé qui ressemble à l'éthylène, mais qui a un cycle benzénique fixé à l'un de ses atomes de carbone. Ce produit a été appelé le Buna S. Ses propriétés sont très proches de celles du caoutchouc naturel ; en fait, grâce au buna S, les forces allemandes n'ont pas eu à souffrir d'une trop grande pénurie de caoutchouc pendant la Deuxième Guerre mondiale. L'Union Soviétique s'est approvisionnée en caoutchouc de la même façon. Les matières premières peuvent être obtenues à partir du charbon ou du pétrole.

Les États-Unis ne commencèrent que tardivement la fabrication de caoutchouc synthétique en quantités commercialisables, sans doute parce qu'ils ne couraient aucun risque de pénurie de caoutchouc jusqu'en 1941. Mais après Pearl Harbour, ce fut la revanche du caoutchouc synthétique. Les États-Unis commencèrent à produire du buna et un autre caoutchouc synthétique appelé *néoprène,* construit à partir du *chloroprène :*



$$CH_2 = \underset{\displaystyle Cl}{\underset{|}{C}} - CH = CH_2$$

Cette molécule, comme on peut le voir, ne diffère de l'isoprène que par la substitution d'un groupement méthyle par un atome de chlore.

Les atomes de chlore liés à la chaîne du polymère à intervalles réguliers confèrent au néoprène une résistance que le caoutchouc naturel n'a pas. Par exemple, il est plus résistant aux solvants organiques tels que l'essence. A l'heure actuelle, on emploie le néoprène de préférence au caoutchouc pour la fabrication des tuyaux de pompes à essence. Dans le domaine des caoutchoucs synthétiques, le néoprène a été le premier à fournir la preuve qu'un produit sortant du tube à essai n'est pas nécessairement une simple imitation de la nature, mais qu'il peut en représenter une amélioration réelle. Ceci est vrai dans bien d'autres domaines.

On produit maintenant des polymères amorphes qui ne ressemblent pas au caoutchouc naturel, mais qui ont des qualités élastiques, et offrent un large éventail de propriétés intéressantes. Étant donné que ce ne sont pas de vrais caoutchoucs, ils sont appelés *élastomères* (abréviation de *polymères élastiques*).

On a découvert, en 1918, le premier élastomère ne ressemblant pas au caoutchouc. C'était un *caoutchouc polysoufré ;* sa molécule était composée de paires d'atomes de carbone, alternant avec quatre atomes de soufre. On donna le nom de Thiokol à cette substance, le préfixe *thio* venant du mot grec « soufre ». L'odeur produite lors de sa fabrication retarda celle-ci pendant très longtemps, mais il fut finalement produit commercialement.

Les élastomères ont été aussi formés à partir de monomères acryliques, de fluorocarbones et de silicones. Là, comme dans presque tous les domaines auxquels il touche, le chimiste organicien travaille à la manière d'un artiste, utilisant des matériaux pour créer des formes nouvelles et améliorer la nature.

Chapitre 12

Les protéines

Les acides aminés

Au début de leurs études sur la matière vivante, les chimistes ont remarqué qu'un groupe de substances se comportait d'une manière particulière. Le fait de chauffer ces substances les changeait de liquides en solides, et non l'inverse. Le blanc d'œuf, une substance du lait (*la caséine*) et un composant du sang (*la globuline*) font partie des substances présentant ces propriétés. Le chimiste français Pierre Joseph Macquer classa en 1777 toutes les substances qui coagulent par chauffage dans une catégorie qu'il appela *albumineuses*, de *l'albumen*, nom que l'encyclopédiste romain Pline avait donné au blanc d'œuf.

Quand les chimistes du XIXe siècle analysèrent les substances albumineuses, ils trouvèrent que ces composés étaient considérablement plus complexes que les autres molécules organiques. En 1839, le chimiste hollandais Gerardus Johannes Mulder établit une formule de base, $C_{40}H_{62}O_{12}N_{10}$, pensant qu'elle était commune à toutes les substances albumineuses. Il croyait que les différents composés albumineux se formaient par addition de petits groupements contenant du soufre ou du phosphore sur cette unité. Mulder appela *protéine* (terme qui lui fut suggéré par le forgeur invétéré de noms, Berzelius) sa formule de base, du mot grec signifiant « de première importance ». Le mot devait indiquer simplement que cette formule centrale était de première importance pour la détermination de la structure des substances albumineuses. Mais, telles que les choses tournèrent, cela se révéla

s'appliquer aux substances proprement dites. Les *protéines,* au fur et à mesure qu'on les découvrait, s'avérèrent rapidement d'une importance fondamentale pour la vie.

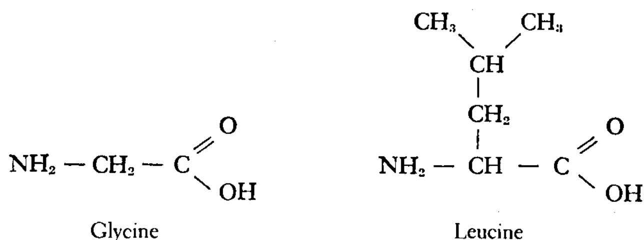
Pendant la décade qui suivit le travail de Mulder, Justus von Liebig établit que les protéines sont encore plus essentielles à la vie que les glucides ou les graisses : elles ne fournissent pas seulement le carbone, l'hydrogène et l'oxygène, mais aussi l'azote, le soufre et souvent le phosphore, qui sont absents des graisses et des glucides.

Les essais de Mulder et des autres pour déterminer la formule empirique complète des protéines étaient condamnés à l'échec à l'époque où ils eurent lieu. La molécule de la protéine, très compliquée, ne pouvait pas être analysée par les méthodes disponibles à ce moment-là. Cependant, une autre approche du problème allait finalement révéler non seulement la composition, mais également la structure des protéines. Les chimistes commençaient à comprendre la nature des unités constitutives des protéines.

En 1820, Henri Braconnot réussit, par chauffage dans de l'acide, à casser la cellulose en unités de glucose (voir chapitre 11), et il décida d'appliquer le même traitement à la gélatine, une substance albumineuse. Le traitement produisait une substance sucrée et cristalline. Contrairement à la première hypothèse de Braconnot, ce produit n'était pas un sucre, mais un composé contenant de l'azote, car on pouvait en obtenir de l'ammoniac (NH_3). Comme on donne par convention, aux substances contenant de l'azote, un nom se terminant par *-ine,* le composé isolé par Braconnot fut appelé *glycine,* du mot grec signifiant « sucré ».

Peu de temps après, Braconnot obtint une substance blanche et cristalline, en chauffant du tissu musculaire en présence d'acide : il l'appela *leucine,* du mot grec signifiant « blanc ».

Finalement, on trouva des ressemblances fondamentales entre la glycine et la leucine, quand leurs formules développées furent établies :

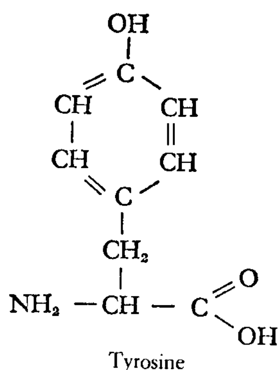


Glycine

Leucine

Chacun de ces composés, comme on peut le voir, possède à une extrémité un groupement aminé (NH_2) et à l'autre extrémité un groupement carboxyle (COOH). Les molécules de ce genre furent appelées *acides aminés,* à cause du groupement carboxyle qui confère des propriétés acides à toutes les molécules qui le contiennent. Les molécules ayant des groupements aminé et carboxyle liés par un seul atome de carbone, comme les deux molécules ci-dessus, sont appelées *acides alpha aminés.*

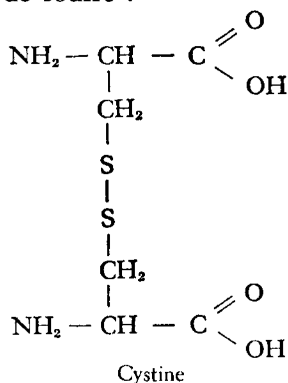
Les années passant, on isola d'autres acides alpha aminés à partir des protéines. Ainsi, Liebig purifia la tyrosine (du mot grec signifiant « fromage ») à partir de la protéine du lait (la caséine, mot provenant de « fromage » en latin).



Tyrosine

Les différences entre ces acides aminés résident entièrement dans la nature des groupements d'atomes attachés à l'atome de carbone reliant les groupements aminé et carboxyle. La glycine, le plus simple de tous les acides aminés, n'a qu'une paire d'atomes d'hydrogène attachés à l'atome de carbone alpha, là où les autres acides aminés possèdent une *chaîne latérale* contenant des atomes de carbone.

Je ne donnerai que la formule de la *cystine*, un des acides aminés que nous rencontrerons plus loin dans ce chapitre. Elle fut découverte en 1899 par le chimiste allemand K.A.H. Mörner ; c'est une molécule à double tête contenant deux atomes de soufre :



Cystine

En réalité, la cystine avait d'abord été isolée par le chimiste anglais William Hyde Wollaston à partir de calculs de la vessie ; d'où son nom dérivé du mot grec pour « vessie ». La contribution de Mörner a consisté à montrer que ce composé vieux d'un siècle était à la fois un composant des protéines et des calculs de la vessie.

La cystine peut être facilement *réduite* (un terme qui, chimiquement, est l'opposé *d'oxydé*), c'est-à-dire qu'elle fixe facilement deux atomes d'hydrogène sur la liaison S-S. La molécule se divise alors en deux moitiés, chacune contenant un groupement -SH (*mercaptan* ou *thiol*). Cette moitié réduite est la *cystéine*, qui est facilement oxydée pour redonner la cystine.

A cause de sa fragilité, le groupement thiol joue un rôle important dans le fonctionnement de nombreuses molécules de protéines. Un équilibre fragile et une capacité à se mouvoir dans un sens ou un autre sous la moindre impulsion sont la marque des produits chimiques les plus

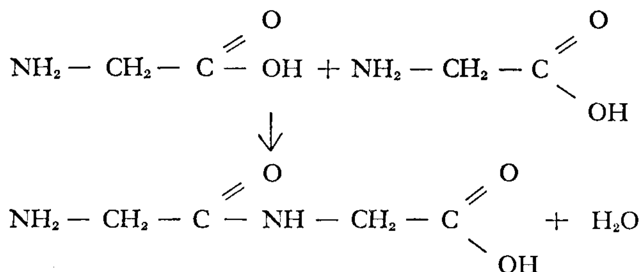
importants du monde vivant ; les membres du groupe des thiols ont permis des combinaisons d'atomes contribuant à cette mobilité.

Actuellement, on a identifié en tout et pour tout dix-neuf acides aminés importants (c'est-à-dire que l'on trouve dans la plupart des protéines). Le dernier a été découvert en 1935 par le chimiste américain William Cumming Rose. Il est peu probable que l'on découvre d'autres acides aminés communs.

LES COLLOÏDES

Vers la fin du XIXe siècle, les biochimistes avaient acquis la certitude que les protéines sont des molécules géantes constituées d'enchaînements d'acides aminés, exactement comme la cellulose est constituée de glucose et le caoutchouc d'unités d'isoprène. Mais il y a une différence importante : alors que la cellulose et le caoutchouc sont faits d'une seule sorte d'unité de construction, une protéine est construite à partir d'un grand nombre d'acides aminés différents. Par conséquent, déterminer la structure d'une protéine pose des problèmes particuliers et subtils.

Le premier problème était de déterminer exactement le mode de liaison des acides aminés dans les chaînes protéiques. Emil Fischer amorça la solution du problème en liant des acides aminés en chaînes, de telle manière que le groupement carboxyle d'un acide aminé soit toujours lié au groupement aminé du suivant. En 1901, il fut le premier à accomplir cette condensation en liant deux molécules de glycine l'une à l'autre par l'élimination d'une molécule d'eau :



C'est la condensation la plus simple. En 1907, Fischer avait synthétisé une chaîne de dix-huit acides aminés, quinze d'entre eux étant des glycines et les trois autres des leucines. Cette dernière molécule ne présentait aucune des propriétés évidentes des protéines, mais Fischer pensa que c'était seulement parce que la chaîne n'était pas assez longue. Il appela les chaînes synthétiques des *peptides,* du mot grec signifiant « digérer », parce qu'il croyait que les protéines se cassaient dans ce type de groupement quand elles étaient digérées. Fischer appela la liaison entre les groupements carboxyle et aminé une *liaison peptidique.*

En 1932, le chimiste allemand Max Bergmann (un élève de Fischer) mit au point une technique pour construire des peptides à partir d'acides aminés variés. En utilisant cette méthode, le biochimiste américain d'origine polonaise, Joseph Stewart Fruton, prépara des peptides qui pouvaient être cassés en des fragments plus petits par des sucs digestifs. Comme on avait de bonnes raisons de croire que les sucs digestifs

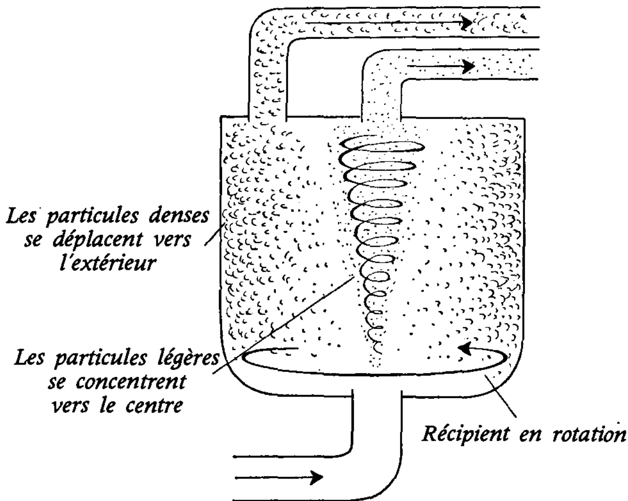
n'*hydrolisaient* (coupaient par addition d'eau) qu'une seule sorte de liaison moléculaire, la liaison entre les acides aminés des peptides synthétiques devait donc être de même nature que celle reliant les acides aminés dans les protéines naturelles. Cette démonstration élimina les derniers doutes concernant la justesse de la théorie de la structure peptidique des protéines établie par Fischer.

Les peptides synthétiques des premières décades du XX[e] siècle étaient très courts et leurs propriétés ne ressemblaient en rien à celles des protéines. Fischer avait synthétisé un peptide long de dix-huit acides aminés, comme je l'ai dit précédemment ; en 1916, le chimiste suisse Emil Abderhalden fit un pas de plus en préparant un peptide de dix-neuf acides aminés, record qu'il détint pendant trente ans. En fait, les chimistes savaient qu'un tel peptide ne représentait qu'un minuscule fragment par rapport à la taille d'une protéine, le poids moléculaire des protéines étant énorme.

Considérons, par exemple, l'hémoglobine, une protéine du sang. L'hémoglobine contient du fer, qui ne constitue que 0,34 % du poids de la molécule. Des expériences chimiques indiquent que la molécule d'hémoglobine possède quatre atomes de fer, de sorte que le poids moléculaire total de l'hémoglobine est de 67 000 environ : quatre atomes de fer d'un poids total de $55{,}85 \times 4$ représentent 0,34 % du poids moléculaire. En conséquence, l'hémoglobine doit contenir environ 550 acides aminés (le poids moléculaire moyen des acides aminés étant de 120 environ). Comparés à ce chiffre, les dix-neuf acides aminés d'Abderhalden semblent mesquins ; et pourtant l'hémoglobine n'est qu'une protéine de taille moyenne.

Les meilleures estimations du poids moléculaire des protéines ont été obtenues en les faisant tourner rapidement dans une *centrifugeuse*, appareil qui, en tournant, pousse les particules du centre vers l'extérieur sous l'action de la force centrifuge (figure 12.1). Quand la force centrifuge est

Figure 12.1. Principe de la centrifugeuse.



plus intense que celle de la gravitation terrestre, les particules en suspension dans le liquide se déplacent vers l'extérieur, loin du centre, plus vite qu'elles ne s'accumuleraient vers le bas sous la force de la pesanteur. Par exemple, les globules rouges se séparent rapidement du plasma dans une centrifugeuse de ce genre, et le lait frais se sépare en deux, la crème grasse et le lait maigre plus dense. Ces séparations sont obtenues très lentement sous l'action de forces de gravitation normales ; la centrifugation les accélère.

Les molécules de protéines, quoique très grandes, ne sont pas assez lourdes pour sortir de la solution sous l'effet de la gravité ; elles ne se séparent pas non plus rapidement dans une centrifugeuse ordinaire. Mais en 1923, le chimiste suédois Theodor Svedberg mit au point une *ultracentrifugeuse* capable de séparer les molécules selon leur poids ; cet engin à grande vitesse tourne à plus de 10 000 tours par seconde et produit des forces centrifuges jusqu'à 900 000 fois plus intenses que la force de gravité à la surface de la Terre. Pour cette contribution à l'étude des suspensions, Svedberg reçut le prix Nobel de chimie en 1926.

Grâce à l'ultracentrifugeuse, les chimistes ont pu mesurer le poids moléculaire de nombreuses protéines d'après leur vitesse de sédimentation (mesurée en svedbergs, en l'honneur du chimiste). On a constaté que les petites protéines ont des poids moléculaires de quelques milliers seulement et ne contiennent vraisemblablement pas plus de cinquante acides aminés (mais en tout cas plus de dix-neuf). D'autres protéines ont des poids moléculaires de l'ordre de la centaine de mille et même du million, ce qui signifie qu'elles sont composées de milliers ou dizaines de milliers d'acides aminés. Le fait de posséder des molécules aussi grandes place les protéines dans une catégorie de substances qui n'ont été étudiées systématiquement que depuis la deuxième moitié du XIXᵉ siècle.

Le chimiste écossais Thomas Graham a été un pionnier en la matière en raison de son intérêt pour la *diffusion,* c'est-à-dire la manière dont les molécules de deux substances mises en contact se mélangent. Il étudia pour commencer la vitesse de diffusion des gaz à travers des petits trous ou dans des tubes très fins. Dès 1831, il pouvait démontrer que la vitesse de diffusion d'un gaz était inversement proportionnelle à la racine carrée de son poids moléculaire (*la loi de Graham*). A ce sujet, on sépare l'uranium 235 de l'uranium 238 en appliquant la loi de Graham.

Pendant les décades qui suivirent, Graham étudia la diffusion des substances dissoutes. Il trouva que les solutions de composés tels que les sels, les sucres ou le sulfate de cuivre passent à travers une paroi de parchemin (contenant probablement des trous submicroscopiques), ce qui n'est pas possible avec des solutions de composés tels que la gomme arabique, la colle et la gélatine. De toute évidence, les molécules géantes de ce dernier groupe de substances ne peuvent pas passer à travers les trous du parchemin.

Graham appela les composés qui pouvaient passer à travers le parchemin (et qui pouvaient être obtenus facilement sous forme cristalline) des *cristalloïdes,* et ceux qui ne le traversaient pas, comme la colle (en grec *kolla*), des *colloïdes.* L'étude des molécules géantes (ou agrégats d'atomes, même quand ceux-ci ne forment pas de molécules distinctes) fut donc connue sous le nom de *chimie des colloïdes.* Comme les protéines, ainsi que d'autres molécules clés du tissu vivant, sont de taille géante, la chimie des colloïdes est d'une importance toute particulière en *biochimie* (l'étude des réactions chimiques ayant lieu dans le tissu vivant).

On peut tirer parti de la taille géante des molécules de protéines de nombreuses façons. Supposons que l'on ait de l'eau pure d'un côté d'une feuille de parchemin, et une solution colloïdale de protéines de l'autre côté. Les molécules de protéines ne peuvent pas passer à travers le parchemin ; de plus, elles bloquent le passage à des molécules d'eau qui, autrement, l'auraient traversé. C'est pour cette raison que l'eau se déplace plus facilement vers la partie colloïdale du système. Le liquide s'accumule du côté de la solution de protéines et y crée une *pression osmotique.*

En 1877, le botaniste allemand Wilhelm Pfeffer montra comment on pouvait mesurer la pression osmotique et déterminer à partir d'elle le poids moléculaire d'une molécule géante. Ce fut la première méthode valable permettant d'estimer la taille de telles molécules.

A nouveau, les solutions de protéines pouvaient être mises dans des sacs faits de *membranes semi-perméables* (des membranes ayant des pores suffisamment larges pour ne laisser passer que les petites molécules, pas les grosses). Si l'on place ces sacs dans de l'eau courante, les petites molécules et les ions passent vers l'extérieur à travers la membrane et sont éliminés, tandis que les grosses molécules de protéines ne sortent pas du sac. Cette méthode, *la dialyse,* est la plus simple pour purifier les protéines en solution.

Contrairement aux petites molécules, les molécules de la taille des colloïdes sont suffisamment grandes pour diffuser la lumière. De plus, la lumière de courte longueur d'onde est diffusée plus efficacement que celle de grande longueur d'onde. Le premier à avoir remarqué cet effet, en 1869, fut le physicien irlandais John Tyndall ; on l'appelle donc *l'effet Tyndall.* La diffusion de la lumière de petite longueur d'onde par des particules de poussière explique la couleur bleue du ciel. En revanche, au coucher du Soleil, la lumière passant à travers une épaisseur plus grande d'atmosphère, rendue particulièrement poussiéreuse par l'activité de la journée, est suffisamment diffusée pour ne laisser passer que le rouge et l'orange, ce qui explique la belle couleur rouge du Soleil couchant.

En passant à travers une solution colloïdale, la lumière se disperse de telle façon qu'on peut voir un cône lumineux quand on se place sur le côté. Les solutions de substances cristallines ne forment pas de cône de lumière quand elles sont éclairées ; elles sont *optiquement claires.* En 1902, le chimiste germano-autrichien Adolf Zsigmondy tira profit de cette observation et inventa un *ultramicroscope* permettant d'examiner les solutions colloïdales à angle droit, les particules individuelles (trop petites pour être visibles au microscope ordinaire) étant perçues comme des points lumineux très brillants. Pour cette invention, il reçut le prix Nobel de chimie en 1925.

Les chimistes des protéines souhaitaient vivement synthétiser *des chaînes polypeptidiques* longues, espérant ainsi produire des protéines. Or, les méthodes de Fischer et Bergmann ne permettaient d'ajouter qu'un seul acide aminé à la fois – procédé qui semblait alors totalement inadéquat. On avait donc besoin d'une technique qui permettrait aux acides aminés de se lier par une sorte de réaction en chaîne, comme celle que Baekeland avait utilisée pour ses hauts polymères plastiques. En 1947, le chimiste israélien E. Katchalski et le chimiste de Harvard, Robert Woodward (celui qui a synthétisé la quinine), réussirent à produire un polypeptide à l'aide d'une réaction en chaîne. Le matériel de départ était un acide aminé légèrement modifié (cette modification s'éliminant pendant la réaction).

En partant de là, ils construisirent des polypeptides synthétiques contenant jusqu'à des centaines, voire des milliers d'acides aminés.

Ces chaînes sont habituellement composées d'une seule sorte d'acide aminé, telle que la glycine ou la tyrosine, et sont donc appelées *polyglycine* ou *polytyrosine*. On peut aussi former un polypeptide ayant une chaîne contenant deux acides aminés différents, en partant d'un mélange de deux acides aminés modifiés. Ces constructions synthétiques ne ressemblent qu'aux protéines les plus simples, par exemple *la fibroïne*, la protéine de la soie.

LES CHAÎNES POLYPEPTIDIQUES

Certaines protéines forment des fibres et sont cristallines comme la cellulose ou le nylon, par exemple la fibroïne, la kératine, protéine du cheveu et de la peau, ou encore le collagène, protéine des tendons et du tissu conjonctif. Le physicien allemand R.O. Herzog a prouvé que ces substances ont un aspect cristallin en montrant qu'elles diffractent les rayons X. Un autre physicien allemand, Rudolf Brill, après analyse de leur diagramme de diffraction, a déterminé l'espacement des atomes dans la chaîne polypeptidique. Utilisant aussi la diffraction des rayons X, le biochimiste anglais William Thomas Astbury et d'autres ont, dans les années trente, obtenu des informations supplémentaires sur la structure de la chaîne. Ils ont réussi à calculer avec une certaine précision la longueur des liaisons entre atomes adjacents et les angles qu'elles forment. Ils ont aussi appris que la chaîne de fibroïne est complètement étirée : ses atomes sont alignés au maximum de ce que permettent les angles des liaisons.

Ce déploiement maximum de la chaîne polypeptidique est l'organisation la plus simple : la *configuration bêta*. Quand un cheveu est étiré, sa molécule de kératine, comme la fibroïne, adopte cette configuration. Si le cheveu est humidifié, il peut s'allonger de trois fois sa longueur initiale. Mais la kératine, à l'état normal, non étirée, est organisée d'une manière plus complexe, la *configuration alpha*.

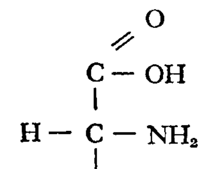
En 1951, Linus Pauling et Robert Brainard Corey, de l'Institut de technologie de Californie, émirent l'hypothèse qu'un polypeptide, dans la conformation alpha, prenait la forme d'une *hélice* (comme un escalier en spirale). Après avoir construit différents modèles pour voir comment la structure pouvait s'organiser, si toutes les liaisons entre les atomes demeuraient dans leur orientation naturelle, sans contrainte, ils conclurent que chaque tour d'hélice devait mesurer 3,6 acides aminés ou 5,4 angströms.

Qu'est-ce qui permet à l'hélice de maintenir sa structure? Pauling suggéra que le responsable est ce que l'on appelle la *liaison hydrogène*. Comme nous l'avons vu lorsqu'un atome d'hydrogène est lié à un atome d'oxygène ou à un atome d'azote, ces derniers retiennent la majeure partie des électrons de liaison, de sorte que l'hydrogène a une légère charge positive et que l'oxygène ou l'azote ont une légère charge négative. Il se trouve que dans l'hélice, un atome d'hydrogène se retrouve périodiquement près d'un atome d'oxygène ou d'azote, situé sur le tour d'hélice immédiatement supérieur ou inférieur. La faible charge positive de l'atome d'hydrogène est attirée par la faible charge négative de son voisin. Cette

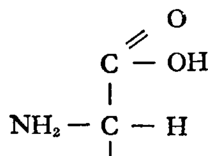
attraction ne représente qu'un vingtième de la force d'une liaison chimique ordinaire, mais elle est suffisamment forte pour maintenir l'hélice en place. Cependant, en exerçant une traction sur la fibre, on déroule facilement l'hélice, et de ce fait on étend la fibre.

Jusqu'ici, nous n'avons considéré que la « colonne vertébrale » de la molécule protéique qui est représentée par ...CCNCCNCCNCCN... Mais les différentes chaînes latérales des acides aminés jouent aussi un rôle important dans la structure de la protéine.

Tous les acides aminés, sauf la glycine, ont au moins un atome de carbone asymétrique, celui qui est situé entre les groupements carboxyle et aminé. Chacun peut donc exister sous deux formes isomériques. Les formules générales des deux isomères sont les suivantes :



Chaîne latérale
Acide aminé D



Chaîne latérale
Acide aminé L

Cependant, il semble tout à fait certain, à partir des données de l'analyse chimique et de celles des rayons X, que les chaînes polypeptidiques ne sont constituées que d'acides aminés de la forme L. Dans cette situation, les chaînes latérales pointent vers l'extérieur, alternativement d'un côté du squelette, puis de l'autre. Une chaîne composée d'un mélange des deux isomères ne serait pas stable, parce que chaque fois qu'un acide aminé L et un acide aminé D se retrouveraient côte à côte, les deux chaînes latérales feraient saillie du même côté, créant de ce fait un entassement et une contrainte sur les liaisons.

Les chaînes latérales jouent un rôle important dans le maintien des liaisons entre chaînes polypeptidiques voisines. Partout où une chaîne latérale chargée négativement est proche d'une chaîne latérale chargée positivement, il se forme une liaison ionique. Les chaînes latérales fournissent aussi des liaisons hydrogènes qui peuvent servir de liens. La cystine, l'acide aminé à deux têtes, peut insérer une de ses séquences amino-carboxyle dans une chaîne et l'autre séquence dans la chaîne voisine. Les deux chaînes sont alors liées ensemble par les deux atomes de soufre de la chaîne latérale (le *pont disulfure*). La liaison entre les chaînes rend compte de la résistance des fibres protéiques. Cela explique la solidité de la toile d'araignée, qui semble si fragile, et le fait que la kératine puisse donner des structures aussi dures que les ongles, les griffes du tigre, les écailles du crocodile et la corne du rhinocéros.

LES PROTÉINES EN SOLUTION

Tout cela décrit fort bien la structure des fibres protéiques. Mais qu'en est-il des protéines en solution ? Quelle structure ont-elles ?

Elles ont certainement une structure définie, mais qui est très fragile. Un chauffage doux, l'agitation d'une solution, l'addition d'un peu d'acide ou d'alcali, ou n'importe quelle autre contrainte du milieu ambiant *dénature*

une protéine en solution, c'est-à-dire que la protéine perd sa capacité d'accomplir ses fonctions normales, et beaucoup de ses propriétés changent. De plus, la dénaturation est généralement irréversible : par exemple, un œuf dur ne peut pas redevenir mou.

Il semble certain que la dénaturation implique la perte de la configuration du squelette peptidique. Quelle est la structure caractéristique détruite ? La diffraction des rayons X ne peut pas nous aider quand les protéines sont en solution, mais on dispose d'autres techniques.

En 1928, par exemple, le physicien indien Chandrasekhara Venkata Raman découvrit que la longueur d'onde de la lumière diffusée par les molécules en solution était modifiée. On pouvait déduire, de la nature des modifications, des informations sur la structure de la molécule. C'est pour cette découverte de l'*effet Raman* que le prix Nobel de physique fut décerné à Raman en 1930. (On se réfère habituellement aux longueurs d'onde modifiées de la lumière sous le nom de *spectre Raman* de la molécule provoquant la diffusion.)

Une autre technique subtile a été mise au point vingt ans plus tard. Elle est basée sur le fait que le noyau atomique a des propriétés magnétiques. Les molécules exposées à un champ magnétique de haute intensité absorbent certaines fréquences de longueur d'onde radio. A partir d'une telle absorption, à laquelle on se réfère en tant que *résonance magnétique nucléaire,* souvent abrégée en RMN, on peut retirer des informations concernant les liaisons entre les atomes. En particulier, les techniques de RMN permettent de localiser la position des petits atomes d'hydrogène dans les molécules, ce que ne peut faire la diffraction des rayons X. Les techniques de RMN ont été mises au point en 1946 par deux équipes travaillant indépendamment : l'une dirigée par E.M. Purcell (qui, plus tard, devait détecter en premier les ondes radio émises par l'atome d'hydrogène neutre dans l'espace ; voir chapitre 2) ; une autre sous la direction du physicien suisse-américain, Felix Bloch. Purcell et Bloch partagèrent le prix Nobel de physique en 1952.

Revenons-en à la dénaturation des protéines en solution : les chimistes américains Paul Mead Doty et Elkan Rogers Blout ont utilisé les techniques de diffusion de la lumière sur des solutions de peptides synthétiques et trouvé qu'ils avaient des structures hélicoïdales. En changeant l'acidité de la solution, Doty et Blout pouvaient transformer les hélices en pelotes statistiques ; en reneutralisant l'acidité, ils pouvaient reformer les hélices. Ils montrèrent aussi que la conversion des hélices en pelotes statistiques réduisait l'amplitude de l'activité optique de la solution. Il était même possible de voir dans quel sens le pas de l'hélice tournait : il tourne dans le sens d'une vis à pas à droite.

Toutes ces découvertes indiquent qu'au cours de la dénaturation d'une protéine, sa structure en hélice est détruite.

LE MORCELLEMENT DE LA MOLÉCULE DE PROTÉINE

Jusqu'ici, nous avons considéré la structure d'une molécule de protéine dans son ensemble, c'est-à-dire la forme générale de la chaîne. Quels sont les détails de sa construction ? Par exemple, combien y a-t-il d'acides aminés de chaque sorte dans une molécule de protéine ?

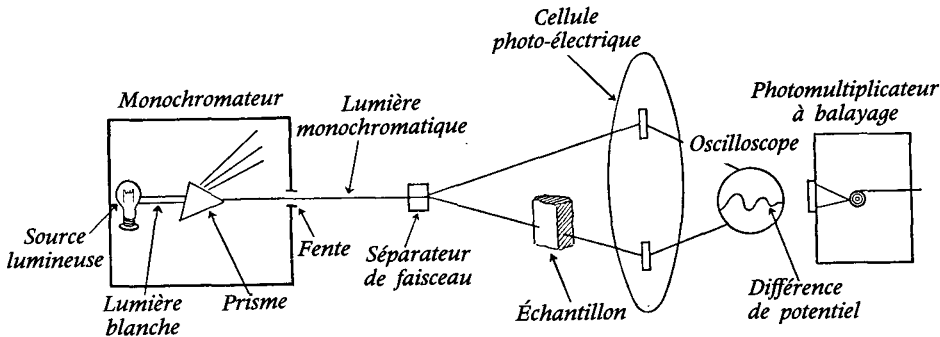
On pourrait casser une molécule protéique en ses acides aminés (en la chauffant en présence d'acide) et déterminer ensuite la proportion de

chaque acide aminé présent dans le mélange. Malheureusement, certains acides aminés se ressemblent tellement chimiquement qu'il est impossible de les séparer nettement par les méthodes chimiques habituelles. Les acides aminés peuvent malgré tout être très bien séparés par chromatographie (voir chapitre 6). En 1941, les biochimistes anglais Archer John Porter Martin et Richard Laurence Millington Synge ont été les premiers à utiliser la chromatographie dans ce but. Ils ont introduit l'amidon en tant que matériel de remplissage des colonnes. En 1948, les biochimistes américains Stanford Moore et William Howard Stein ont porté à un haut degré d'efficacité la chromatographie des acides aminés sur colonne d'amidon, ce qui leur a valu le prix Nobel de chimie en 1972.

Lorsque le mélange d'acides aminés en solution a été déposé sur la colonne d'amidon, et que tous les acides aminés se sont liés aux particules, ils sont lentement entraînés vers le bas par un nouveau solvant. Chaque acide aminé descend le long de la colonne avec sa vitesse caractéristique. Au fur et à mesure que chacun d'entre eux émerge séparément au bas de la colonne, on récolte les gouttes de solution dans un récipient. La solution de chaque récipient est alors traitée par un réactif qui transforme l'acide aminé en un produit coloré. L'intensité de la couleur dépend de la quantité d'acide aminé présente dans la solution. L'intensité de cette coloration est déterminée par un instrument, le *spectrophotomètre,* qui sert à mesurer l'intensité d'absorption de la lumière d'une longueur d'onde donnée (figure 12.2).

A ce sujet, les spectrophotomètres sont aussi utilisés pour d'autres analyses chimiques. Si on envoie à travers une solution de la lumière de longueur d'onde qui augmente petit à petit, la valeur de l'absorption change sans heurts, atteignant un maximum pour certaines longueurs d'onde et tombant à un minimum pour d'autres. Le résultat est un *spectre d'absorption.* Un groupement atomique donné a des pics d'absorption caractéristiques. C'est spécialement vrai dans la région des infrarouges, comme cela a été démontré en premier par le physicien américain William

Figure 12.2. Le spectrophotomètre. Le faisceau lumineux est divisé en deux : un faisceau passe à travers l'échantillon à analyser, et l'autre va directement à la cellule photo-électrique. Étant donné que le faisceau qui passe par l'échantillon est partiellement absorbé, il libère moins d'électrons dans la cellule photo-électrique que le faisceau de référence qui n'est pas absorbé. Cela crée une différence de potentiel qui mesure l'absorption de la lumière par l'échantillon.

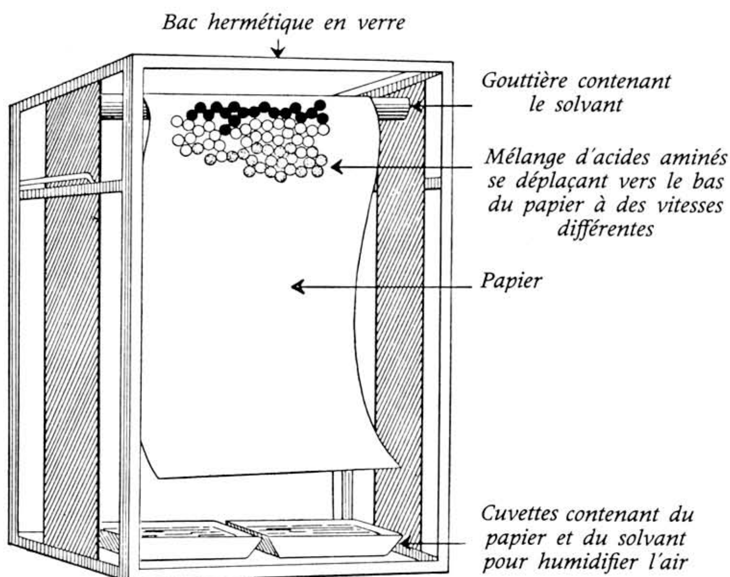


Weber Coblentz au tout début du siècle. Ses instruments étaient alors trop grossiers pour que la technique soit applicable ; mais depuis la Deuxième Guerre mondiale, des *spectrophotomètres à infrarouge*, conçus pour balayer automatiquement le spectre de deux à quarante microns, et pouvant enregistrer les résultats, sont de plus en plus utilisés pour analyser la structure de composés complexes. Les *méthodes optiques* d'analyse chimique comprenant l'absorption de longueurs d'ondes radio et de la lumière, la diffusion de la lumière, etc., sont très raffinées et ne sont pas destructrices. Autrement dit, l'échantillon survit à l'inspection. Ces techniques remplacent complètement les méthodes classiques de Liebig, Dumas et Pregl qui ont été mentionnées dans le chapitre précédent.

Le dosage des acides aminés par la chromatographie sur gel d'amidon est tout à fait satisfaisant ; mais au moment où cette méthode a été développée, Martin et Synge avaient déjà mis au point une méthode de chromatographie plus simple : la *chromatographie sur papier* (figure 12.3).

Les acides aminés sont séparés sur une feuille de papier-filtre (un papier absorbant fait de cellulose particulièrement pure). Une goutte ou deux d'un mélange d'acides aminés sont déposées près d'un coin de la feuille, et le bord de la feuille est ensuite trempé dans un solvant tel que l'alcool butylique. Le solvant imprègne lentement le papier par capillarité (trempez le coin d'un buvard dans l'eau, et voyez par vous-même ce qui se passe). Le solvant capte les molécules de la goutte déposée et les entraîne le long du papier. Comme dans la chromatographie sur colonne, chaque acide aminé se déplace le long du papier avec sa vitesse caractéristique. Au bout d'un moment, les acides aminés du mélange se répartissent en une série de taches sur la feuille. Certaines de ces taches contiennent deux ou trois acides aminés. Pour les séparer, le papier-filtre, après avoir été séché, est

Figure 12.3. La chromatographie sur papier.



tourné de 90° par rapport à sa position première, et le nouveau bord est trempé dans un deuxième solvant qui séparera les composants en taches indépendantes. Enfin, la feuille entière, après avoir été séchée à nouveau, est imprégnée de réactifs qui révèlent les acides aminés sous l'aspect de taches colorées et foncées. La vision est spectaculaire : tous les acides aminés, mélangés au départ en une seule solution, sont alors étalés sur toute la longueur et toute la largeur du papier en une mosaïque de taches colorées. Les biochimistes expérimentés peuvent identifier chaque acide aminé par son emplacement, et peuvent donc lire la composition de la protéine originale presque en un clin d'œil. En dissolvant une tache, ils peuvent même mesurer la quantité de cet acide aminé dans la protéine. Pour la mise au point de cette technique, le prix Nobel de chimie fut attribué à Martin et Synge en 1952.

Martin, en collaboration avec A.T. James, a appliqué les principes de cette technique à la séparation des gaz en 1953. Des mélanges de gaz ou de vapeurs peuvent être passés à travers un solvant liquide, ou même un solide adsorbant, au moyen d'un courant de *gaz entraîneur* inerte, tel que l'hélium ou l'azote. On fait passer le mélange à travers la colonne, et il émerge à l'autre extrémité, séparé en ses composants. La *chromatographie gazeuse* est particulièrement utile, car on obtient une grande vitesse de séparation et on peut détecter les traces d'impuretés avec une grande précision.

L'analyse chromatographique a apporté des évaluations précises de la composition de différentes protéines en acides aminés. Ainsi, on a déterminé que la molécule d'une protéine du sang appelée la *sérum albumine* contient 15 glycines, 45 valines, 58 leucines, 9 isoleucines, 31 prolines, 33 phényl alanines, 18 tyrosines, 1 tryptophane, 22 sérines, 27 thréonines, 16 cystines, 4 cystéines, 6 méthionines, 25 arginines, 16 histidines, 58 lysines, 46 acides aspartiques et 80 acides glutamiques, un total de 526 acides aminés de 18 types différents, associés dans une protéine ayant un poids moléculaire de 69 000 environ (en plus de ces 18 acides aminés, il y en a un autre très courant : l'alanine).

Le biochimiste américain d'origine allemande, Erwin Brand, a proposé un système de symboles pour les acides aminés, qui est maintenant adopté. Pour éviter des confusions avec les symboles des éléments, il a représenté chaque acide aminé par les trois premières lettres de son nom, au lieu de l'initiale seulement. Il y a quelques exceptions : la cystine est symbolisée par CyS, pour montrer que ses deux moitiés sont généralement incorporées dans deux chaînes différentes ; la cystéine par CySH, pour la différencier de la cystine ; et l'isoleucine par Ileu plutôt que par Iso, car *iso* est le préfixe de nombreux termes chimiques.

Dans le langage symbolique, on peut écrire la formule de la sérum albumine comme suit : $Gly_{15}Val_{45}Leu_{58}Ileu_{9}Pro_{31}Phe_{33}Tyr_{18}Try_{1}Ser_{22}$-$Thr_{27}$-$CyS_{32}CySH_{4}Met_{6}Arg_{25}His_{16}Lys_{58}Asp_{46}Glu_{80}$; ce qui, il faut l'admettre, est plus concis, mais imprononçable rapidement.

L'ANALYSE DE LA CHAÎNE PEPTIDIQUE

La découverte de la formule empirique d'une protéine ne représentait que la moitié de la bataille, en fait beaucoup moins que la moitié. Il restait – tâche bien plus difficile – à déchiffrer la structure de la molécule d'une

protéine. Il y avait toutes les raisons de penser que les propriétés de chaque protéine dépendent de l'ordre dans lequel les acides aminés sont arrangés dans la chaîne moléculaire. Cette hypothèse confronte le biochimiste à un problème grave. Le nombre de combinaisons possibles dans lesquelles dix-neuf acides aminés peuvent être arrangés dans une chaîne (même en supposant qu'un seul acide aminé de chaque sorte soit utilisé) est de près de 120 millions de milliards. Si vous avez du mal à croire cela, essayez de multiplier 19 fois 18 par 17 fois 16, etc. C'est la façon dont on calcule le nombre d'arrangements possibles. Si vous ne faites pas confiance à l'arithmétique, prenez dix-neuf pions, numérotez-les de 1 à 19, et voyez de combien de façons différentes vous pouvez les arranger. Je suis sûr que vous ne jouerez pas longtemps.

Quand on a une protéine de la taille de la sérum albumine, qui est composée de plus de 500 acides aminés, le nombre d'arrangements possibles devient de l'ordre de 10^{600} (c'est-à-dire 1 suivi de 600 zéros). C'est un chiffre complètement extravagant, bien supérieur au nombre de particules subatomiques contenues dans tout l'Univers – ou bien plus que l'Univers ne pourrait en contenir s'il était rempli de telles particules sans espace libre.

Néanmoins, bien qu'il puisse sembler sans espoir de trouver la bonne combinaison d'acides aminés parmi toutes celles qui sont possibles dans une molécule de sérum albumine, ce genre de problème a en fait été abordé et résolu.

En 1945, le biochimiste anglais Frederick Sanger décida de déterminer l'ordre des acides aminés dans une chaîne peptidique. Il commença par identifier l'acide aminé situé à l'extrémité amine terminale de la chaîne.

Il va de soi que le groupement aminé de cet acide aminé terminal (appelé *acide aminé N terminal*) est libre, c'est-à-dire qu'il n'est lié à aucun autre acide aminé. Sanger a utilisé un réactif qui, en se combinant à un groupement aminé libre, et non pas à une amine liée à un groupement carboxyle, produit un dérivé DNP (*dinitrophényle*) de la chaîne peptidique. Il pouvait marquer l'acide N terminal avec le DNP, et comme la liaison qui maintient cette association est plus forte que la liaison reliant les acides aminés dans une chaîne, il pouvait casser la chaîne en ses acides aminés individuels, et isoler l'acide aminé marqué par le DNP. Le groupement DNP ayant une couleur jaune, l'acide aminé ainsi marqué par le DNP se révèle par une tache jaune lors d'une chromatographie sur papier.

Sanger réussit donc à séparer et identifier l'acide aminé de l'extrémité aminée de la chaîne peptidique. De façon similaire, il identifia l'acide aminé à l'autre extrémité de la chaîne, celui ayant un groupement carboxyle libre, *l'acide aminé C terminal*. Il put aussi, dans certains cas, décrocher quelques acides aminés un par un et identifier la *séquence terminale* d'une chaîne peptidique.

Ensuite, Sanger s'attaqua à la chaîne peptidique entière. Il travailla sur *l'insuline*, une protéine qui a une fonction importante dans l'organisme et, de plus, est assez petite pour une protéine, avec un poids moléculaire de 6 000 sous sa forme la plus simple. Le traitement par le DNP a montré que cette molécule est constituée de deux chaînes peptidiques, car elle contient deux acides aminés N terminaux différents. Les deux chaînes sont reliées par des molécules de cystine. Après avoir rompu la liaison entre les deux atomes de soufre de la cystine par un traitement chimique, Sanger sépara la molécule d'insuline en ses deux chaînes peptidiques, chacune

étant intacte. L'une des chaînes avait une glycine comme acide aminé N terminal (la chaîne G), et l'autre une phénylalanine N terminale (la chaîne P). On peut maintenant travailler sur chacune séparément.

Sanger et un collaborateur, Hans Tuppy, dégradèrent d'abord les chaînes en leurs acides aminés individuels et identifièrent les vingt et un acides aminés qui composent la chaîne G et les trente acides aminés composant la chaîne P. Ensuite, pour étudier une partie de la séquence, ils dégradèrent les chaînes, non plus en acides aminés individuels, mais en fragments de deux ou trois acides aminés. Cette tâche se faisait soit par hydrolyse partielle, en ne cassant que les liaisons les plus fragiles de la chaîne, soit en attaquant l'insuline au moyen de substances digestives qui ne rompent que certaines liaisons entre les acides aminés, et laissent les autres intactes.

Par ce procédé, Sanger et Tuppy coupèrent chacune des chaînes en de nombreux morceaux. Par exemple, la chaîne P donna 48 fragments différents, dont 22 sont composés de deux acides aminés (*dipeptides*), 14 de trois et 12 de plus de trois.

Tous ces petits peptides étaient séparés, puis scindés en leurs acides aminés individuels, qui étaient ensuite identifiés par chromatographie sur papier. Il restait aux chercheurs à déterminer l'ordre des acides aminés dans ces fragments. Supposons que l'un d'eux soit constitué du dipeptide valine-isoleucine. La question qui se pose est la suivante : l'ordre est-il Val-Ileu ou Ileu-Val ? En d'autres termes, l'acide aminé N terminal est-il la valine ou l'isoleucine ? Le groupement aminé, et par conséquent l'unité N terminale de la chaîne, est considéré par convention comme étant l'extrémité située à gauche. Ici, le marquage par le DNP permet d'obtenir la réponse. Si on le trouve sur la valine, celle-ci est l'acide aminé N terminal, et l'organisation du dipeptide est alors Val-Ileu. Si on le retrouve sur l'isoleucine, l'ordre est alors Ileu-Val.

L'arrangement dans un fragment constitué de trois acides aminés a pu aussi être élucidé. Supposons que les composants soient la leucine, la valine et l'acide glutamique. Le test DNP identifie l'acide aminé N terminal. Si c'est, par exemple, la leucine, l'ordre est soit Leu-Val-Glu, soit Leu-Glu-Val. Ces deux combinaisons étaient synthétisées et déposées sur un chromatogramme pour voir celle qui donnait la même tache sur le papier que le fragment étudié.

Les peptides de plus de trois acides aminés devaient être cassés en des fragments plus petits pour l'analyse.

Après avoir ainsi déterminé la structure de tous les fragments obtenus à partir de l'insuline, il s'agissait de réunir les morceaux dans le bon ordre pour reconstituer la chaîne, à la manière d'un puzzle. On possédait quelques informations. Par exemple, on savait que la chaîne G ne contient qu'un seul acide aminé alanine. Dans le mélange de peptides obtenus lors de la rupture des chaînes G, l'alanine a été retrouvée dans deux combinaisons : alanine-sérine et cystine-alanine. Donc, dans la chaîne G intacte, l'ordre doit être CyS-Ala-Ser.

Au moyen de tels indices, Sanger et Tuppy ont progressivement reconstitué le puzzle. Il leur a fallu deux ans pour identifier tous les fragments et les arranger en une séquence satisfaisante ; mais, en 1952, ils avaient déterminé l'arrangement exact de tous les acides aminés dans les chaînes G et P. Ils établirent aussi la manière dont les deux chaînes sont reliées. En 1953, on a pu enfin annoncer leur triomphe concernant le déchiffrage de la structure de l'insuline. La structure complète d'une

importante protéine venait d'être élucidée pour la première fois. Pour cet exploit, le prix Nobel de chimie fut décerné à Sanger en 1958.

Les biochimistes ont immédiatement adopté la méthode de Sanger pour déterminer la structure d'autres molécules protéiques. Ils y parvinrent, en 1959, avec la ribonucléase, une protéine formée d'une seule chaîne polypeptidique de 124 acides aminés ; puis, en 1960, avec la protéine constituant la sous-unité du virus de la mosaïque du tabac, composée de 158 acides aminés. En 1964, la trypsine, une protéine de 223 acides aminés, était déchiffrée. Dès 1967, la technique était automatisée. Le biochimiste australo-suédois Pehr Edman conçut un *séquenceur* ou *séquanator,* qui pouvait travailler sur cinq milligrammes de protéine pure, en détachant et en identifiant les acides aminés un par un. Soixante acides aminés de la myoglobine furent ainsi identifiés en quatre jours.

Des peptides encore plus longs ont été étudiés dans tous leurs détails ; et il était évident que la structure de n'importe quelle protéine purifiée, même de grande taille, pourrait être déterminée dans les années quatre-vingts. Il suffisait de s'atteler à la tâche.

De telles analyses ont, en général, montré que la plupart des protéines ont des acides aminés variés bien représentés le long de leur chaîne. Seules certaines protéines fibreuses simples, comme celles trouvées dans la soie ou dans les tendons, ont une composition fortement déséquilibrée en faveur de deux ou trois acides aminés.

Les acides aminés individuels sont alignés sans ordre apparent dans ces protéines constituées de dix-neuf acides aminés ; on ne peut pas y repérer facilement de répétitions périodiques. Au lieu de cela, l'organisation des acides aminés permet le repliement de la chaîne grâce à la formation localisée de liaisons hydrogènes et la délimitation par les différentes chaînes latérales d'un espace où l'organisation adéquate des groupements atomiques et des charges électriques permet à la protéine d'accomplir sa tâche.

LES PROTÉINES SYNTHÉTIQUES

Une fois déterminé l'ordre des acides aminés dans la chaîne peptidique, on pouvait essayer d'assembler les acides aminés dans cet ordre. On commença bien sûr par un cas simple. L'oxytocine fut la première protéine synthétisée en laboratoire, cette hormone ayant d'importantes fonctions dans l'organisme. C'est une protéine extrêmement petite : elle n'est constituée que de huit acides aminés. En 1953, le biochimiste américain Vincent du Vigneaud réussit à synthétiser une chaîne peptidique exactement semblable à celle que l'on estimait représenter la molécule d'oxytocine. Le peptide synthétique avait bien toutes les propriétés de l'hormone naturelle, et le prix Nobel de chimie fut décerné à du Vigneaud en 1955.

Dans les années qui suivirent, des molécules de protéines plus compliquées furent synthétisées ; pour synthétiser une molécule donnée, ayant des acides aminés ordonnés d'une façon définie, les perles devaient être, pour ainsi dire, enfilées une à une. C'était aussi difficile, dans les années cinquante, que ça l'avait été un demi-siècle auparavant, du temps de Fischer. Chaque fois qu'un acide aminé donné devait être couplé à la chaîne, il fallait séparer le nouveau composé de tout le reste par des

méthodes très fastidieuses, et il fallait tout recommencer chaque fois que l'on ajoutait un nouvel acide aminé à la chaîne. A chaque étape, une bonne partie du matériel était perdue dans des réactions secondaires, et seules de petites quantités de chaînes, même simples, était obtenues.

Cependant au début de 1959, une équipe, sous la direction du biochimiste américain Robert Bruce Merrifield, inventa une nouvelle méthode. On lie un acide aminé, celui du début de la chaîne désirée, à des billes de résine de polystyrène. Ces billes sont insolubles dans le liquide utilisé et peuvent être séparées de tout le reste par simple filtration. Une nouvelle solution contenant l'acide aminé suivant peut alors être introduite, et cet acide aminé se lie au premier. On filtre à nouveau, puis on ajoute un nouvel acide aminé. Les étapes entre les ajouts sont si simples et si rapides qu'elles peuvent être automatisées avec très peu de pertes. En 1965, on synthétisa ainsi la molécule d'insuline ; en 1969, c'était le tour de la ribonucléase, avec sa chaîne encore plus longue de 124 acides aminés. Puis en 1970, le biochimiste américain d'origine chinoise Cho Hao Li synthétisa l'hormone de croissance longue de 188 acides aminés. En principe, n'importe quelle protéine peut maintenant être synthétisée ; il suffit de s'en donner la peine.

LA FORME DE LA MOLÉCULE PROTÉIQUE

On savait désormais que la molécule protéique était un enchaînement d'acides aminés ; il restait à s'en faire une idée plus précise. De quelle manière exactement la chaîne d'acides aminés se replie-t-elle ? Quelle est la forme exacte de la molécule protéique ?

Le chimiste austro-anglais Max Ferdinand Perutz et son collègue anglais John Cowdery Kendrew s'attelèrent à ce problème. Perutz se consacra à l'hémoglobine, la protéine du sang qui transporte l'oxygène et contient quelque douze mille atomes. Kendrew se chargea de la myoglobine, une protéine du muscle ayant une fonction similaire à celle de l'hémoglobine, mais quatre fois plus petite. Pour mener à bien leurs recherches, ils se servirent de la diffraction des rayons X.

Perutz combina astucieusement les molécules de protéines à des atomes lourds, tels que ceux de l'or et du mercure, qui sont particulièrement efficaces dans la diffraction des rayons X. Il obtint ainsi des indications qui lui permirent de retracer plus précisément la structure de la molécule, ce qu'il n'aurait pu faire sans la présence de ces atomes lourds. Dès 1969, la structure de la myoglobine, et l'année suivante, celle de l'hémoglobine furent mises en place. Il devint possible de construire des modèles tridimensionnels où la place de chaque atome isolé pouvait être déterminée avec une bonne certitude. Dans les deux cas, l'hélice était la base de la structure de la protéine. Ces travaux valurent aux deux chercheurs de partager le prix Nobel de chimie en 1962.

On peut penser que les structures tridimensionnelles établies par les techniques de Perutz et Kendrew sont, en définitive, déterminées par la nature de l'enchaînement des acides aminés. Celui-ci présente des points de repliement naturels ; et quand il y a pliure, certaines liaisons ont inévitablement lieu, qui maintiennent correctement le repliement. Il est possible de déterminer ces repliements et ces liaisons en mesurant toutes les distances des liaisons interatomiques et les angles que les liaisons font

entre elles, mais un tel travail est fastidieux. Là aussi, on a fait appel aux ordinateurs qui ont non seulement fait les calculs, mais aussi projeté les résultats sur écran.

La liste des molécules protéiques dont la forme est connue dans les détails tridimensionnels croît rapidement. La structure tridimensionnelle de l'insuline, qui fut le point de départ de ces nouvelles incursions dans la biologie moléculaire, fut déterminée par la biochimiste anglaise Dorothy Crowfoot Hodgkin en 1969.

Les enzymes

L'immense variété et la complexité des protéines sont bénéfiques. Les protéines remplissent une multitude de fonctions dans les organismes vivants.

Une de ces fonctions est de pourvoir à la structure de la charpente du corps. Les protéines fibreuses jouent ce rôle de charpente chez les animaux complexes, tout comme la cellulose chez les plantes. Les araignées tissent des fils arachnéens et les larves d'insectes les fils de leur cocon avec une protéine, la kératine. Les écailles de poissons et de reptiles sont principalement faites de kératine. Les cheveux, les plumes, les cornes, les sabots, les griffes et les ongles sont tout simplement des écailles modifiées, qui contiennent aussi de la kératine. Les tissus internes de soutien – les cartilages, les ligaments, les tendons ainsi que le squelette organique des os – sont largement constitués de molécules protéiques, telles que le collagène et l'élastine. Les muscles sont faits d'une protéine fibreuse complexe appelée l'*actomyosine.*

Dans tous les cas, les fibres protéiques sont plus qu'un équivalent de la cellulose. Elles en sont une amélioration ; elles sont plus solides et plus flexibles. La cellulose suffit à soutenir une plante à laquelle on ne demande rien d'autre que de se balancer au vent. Mais les fibres protéiques sont conçues pour permettre la flexion des différentes parties du corps, les mouvements, etc.

Cependant, les fibres, dans leur forme aussi bien que dans leur fonction, font partie des protéines les plus simples. La plupart des autres protéines ont des tâches plus complexes à accomplir.

Pour maintenir la vie sous tous ses aspects, de nombreuses réactions chimiques ont lieu dans l'organisme. Elles se déroulent rapidement et sont très variées. Chaque réaction s'enchaîne avec toutes les autres, car le déroulement normal de la vie ne dépend pas d'une seule réaction, mais de l'ensemble de celles-ci. De plus, toutes les réactions doivent avoir lieu dans de bonnes conditions – à températures et pressions modérées, et sans produits chimiques drastiques. Les réactions doivent être contrôlées de façon souple mais rigoureuse, et constamment ajustées aux modifications de l'environnement et aux besoins changeants de l'organisme. Le ralentissement ou l'accélération injustifiés d'une seule des milliers de réactions peut dérégler l'organisme plus ou moins gravement.

Tout cela est rendu possible grâce aux protéines.

LA CATALYSE

Vers la fin du XVIIIe siècle, les chimistes, sous la conduite de Lavoisier, ont étudié les réactions quantitativement ; en particulier, ils ont mesuré les vitesses des réactions. Ils se sont aperçus rapidement que des changements mineurs du milieu pouvaient causer des modifications considérables des vitesses des réactions. Par exemple, Kirchhoff découvrit que l'amidon pouvait être converti en sucre en présence d'acide, et il remarqua que, tandis que l'acide augmente fortement la vitesse de la réaction, il n'est pas consommé pendant le processus. On découvrit rapidement d'autres cas semblables. Le chimiste allemand Johann Wolfgang Döbereiner trouva que le platine finement divisé (appelé *noir de platine*) facilitait l'association de l'hydrogène avec l'oxygène pour former de l'eau ; une réaction qui, sans cette aide, ne se faisait qu'à forte température. Döbereiner conçut même une lampe auto-allumante dans laquelle un jet d'hydrogène, jouant sur une surface recouverte de noir de platine, prenait feu.

Comme les « réactions accélérées » consistaient en général à dégrader des substances complexes en des substances plus simples, Berzelius appela le phénomène *catalyse* (du mot grec signifiant « dissolution »). Le noir de platine fut donc appelé catalyseur de la réaction d'association de l'hydrogène avec l'oxygène, et l'acide, catalyseur de l'hydrolyse de l'amidon en glucose.

La catalyse s'est avérée de la plus grande importance dans l'industrie. Par exemple, la meilleure façon de fabriquer de l'acide sulfurique (le produit de chimie organique le plus important après l'air, l'eau et peut-être le sel) implique l'oxydation du soufre, d'abord en anhydride sulfureux (SO_2), puis en anhydride sulfurique (SO_3). L'étape de conversion de l'anhydride sulfureux en anhydride sulfurique ne se produirait qu'à la vitesse d'un escargot sans l'aide d'un catalyseur comme le noir de platine. Le nickel finement divisé (qui a remplacé le noir de platine dans la plupart des cas, parce qu'il est moins cher), ainsi que des composés comme le chromite de cuivre, le pentoxyde de vanadium, l'oxyde ferrique et le bioxyde de manganèse, sont des catalyseurs importants. En fait, une grande partie du succès d'un procédé dans l'industrie chimique dépend de la découverte du bon catalyseur pour la réaction impliquée. La découverte d'un nouveau type de catalyseur par Ziegler révolutionna la production des polymères.

Comment une substance peut-elle être présente à de très faibles concentrations, et pourtant provoquer une réaction importante, sans être elle-même modifiée ?

Certains catalyseurs prennent en fait part à la réaction, mais d'une façon circulaire, de sorte qu'ils sont continuellement reformés. Un exemple en est le pentoxyde de vanadium (V_2O_5), quand il catalyse le passage de l'anhydride sulfureux à l'anhydride sulfurique. Le pentoxyde de vanadium donne un de ses atomes d'oxygène au SO_2, formant SO_3 et devenant lui-même l'oxyde de vanadium (V_2O_4). Mais l'oxyde de vanadium réagit rapidement avec l'oxygène de l'air et redevient V_2O_5. Le pentoxyde de vanadium agit donc comme un intermédiaire, transmettant un atome d'oxygène à l'anhydride sulfureux, et en prenant un autre à l'air, le transmettant à l'anhydride sulfureux, etc. Le processus est si rapide qu'une petite quantité de pentoxyde de vanadium suffit à convertir de grandes

quantités d'anhydride sulfureux ; à la fin, le pentoxyde de vanadium apparaît inchangé.

En 1902, le chimiste allemand George Lunge émit l'hypothèse que ce processus expliquait la catalyse en général. En 1916, Irving Langmuir fit un pas de plus et proposa une explication de l'activité catalytique d'une substance telle que le platine, qui n'est pas réactive, et que l'on ne peut pas s'attendre à voir s'engager dans des réactions chimiques ordinaires. Langmuir suggéra que les liaisons de valence en excès à la surface du platine métallique capturent les molécules d'hydrogène et d'oxygène. Tandis que ces dernières sont emprisonnées et très proches les unes des autres et de la surface du platine, elles deviennent bien plus aptes à s'associer pour former de l'eau que dans leur état normal de molécules gazeuses libres. Une fois la molécule d'eau formée, elle est déplacée de la surface du platine par des molécules d'oxygène et d'hydrogène. Le procédé comprend la capture de l'hydrogène et de l'oxygène, leur association pour former de l'eau, la libération de l'eau, la capture de plus d'hydrogène et d'oxygène, et la formation de plus d'eau, et peut continuer ainsi indéfiniment.

Ce processus est appelé *la catalyse de surface.* Naturellement, plus le métal est finement divisé, plus la surface offerte par rapport à la masse du métal sera grande, et plus la catalyse sera efficace. Mais si une substance étrangère s'attache fermement aux liaisons de surface du platine, elle *empoisonne* le catalyseur.

Les catalyseurs de surface sont plus ou moins sélectifs, ou *spécifiques.* Certains absorbent facilement les molécules d'hydrogène et catalysent des réactions impliquant l'hydrogène ; d'autres absorbent facilement des molécules d'eau, et catalysent des condensations ou des hydrolyses, etc.

On peut utiliser pour d'autres usages que la catalyse la capacité, très répandue, de certaines surfaces à fixer des couches de molécules (*adsorption*). Le dioxyde de silicium préparé sous forme spongieuse (*gel de silice*) adsorbe de grandes quantités d'eau. Il agit comme *desséchant,* conservant un faible taux d'humidité. Par exemple, on le place dans les emballages d'équipements électroniques qui, sans cela, souffriraient d'une forte humidité.

Le charbon finement divisé (*charbon activé*) adsorbe facilement les molécules organiques ; plus la molécule organique est grande, plus elle est adsorbée. Le charbon activé peut être utilisé pour décolorer les solutions, car il adsorbe les impuretés colorées (généralement de grand poids moléculaire), laissant la substance désirée (habituellement incolore et de petit poids moléculaire).

Le charbon actif est aussi utilisé dans les masques à gaz, une utilisation pressentie par le médecin anglais John Stenhouse, qui prépara le premier filtre à air à charbon actif en 1853. L'oxygène et l'azote de l'air passent à travers cette masse sans être affectés, mais les molécules du gaz empoisonné, qui sont relativement grosses, sont adsorbées.

LA FERMENTATION

Le monde organique a aussi ses catalyseurs. En réalité, certains sont connus depuis des milliers d'années, mais pas sous ce nom. Ils sont aussi vieux que le pain et le vin.

La pâte à pain, laissée telle et protégée des contaminations de l'influence extérieure, ne monte pas. Ajoutez un peu de *levain* (du mot latin signifiant « lever »), et des bulles apparaissent, qui lèvent et allègent la pâte.

La levure (autre terme pour « levain ») accélère aussi la transformation des jus de fruits et du grain en alcool. Ici encore, la transformation comprend la formation de bulles ; c'est pourquoi le processus est appelé *fermentation,* du mot latin signifiant « bouillir ». On appelle souvent les préparations de levures des *ferments.*

Ce n'est qu'au XVII[e] siècle que la nature du levain a été découverte. En 1680, pour la première fois, le chercheur hollandais Anton Van Leeuwenhoek vit des cellules de levure. Pour cela, il utilisa un instrument qui devait révolutionner la biologie, le *microscope,* dont le principe repose sur la courbure et la convergence de la lumière obtenue à l'aide de lentilles. Les instruments opérant avec des associations de lentilles (*les microscopes composés*) furent conçus dès 1590 par le fabricant de lunettes hollandais Zacharias Janssen. Les premiers microscopes étaient utiles dans leur conception, mais leurs lentilles étaient si mal fabriquées que les objets agrandis formaient des taches floues. Van Leeuwenhoek fabriqua des lentilles minuscules mais parfaites qui agrandissaient jusqu'à deux cents fois, en donnant une image tout à fait nette. Il utilisa des lentilles simples (*microscope simple*).

Avec le temps, l'utilisation de bonnes lentilles groupées s'est répandue (car un microscope composé est, potentiellement du moins, bien plus puissant qu'un microscope simple) ; le monde microscopique s'ouvrait. Un siècle et demi après Leeuwenhoek, un physicien français, Charles Cagniard de La Tour, muni d'un bon microscope composé, étudia les levures avec beaucoup d'attention, et réussit à les surprendre en train de se reproduire. Les petites taches étaient vivantes. En 1850, la levure devint un sujet d'étude de première importance.

L'industrie française du vin était alors en difficulté. Le vin, en vieillissant, devenait aigre, imbuvable, et des millions de francs étaient perdus. Le problème fut soumis au jeune doyen de la faculté des sciences de l'université de Lille. Ce jeune doyen était Pasteur en personne, qui s'était déjà fait une grande réputation pour avoir séparé le premier des isomères optiques en laboratoire.

Pasteur étudia les cellules de levure du vin au microscope, et découvrit que les cellules étaient de types variés. Tous les vins contenaient des levures qui permettaient la fermentation, mais les vins qui devenaient aigres contenaient en plus un autre type de levure. Pasteur pensa que le phénomène d'aigreur commençait une fois la fermentation terminée. Puisque l'on n'avait plus besoin des levures une fois la fermentation nécessaire accomplie, pourquoi ne pas s'en débarrasser à ce moment-là pour éviter que le mauvais type de levure ne cause des problèmes ?

Il suggéra alors aux industriels du vin, horrifiés, que le vin soit chauffé légèrement après la fermentation, pour tuer toutes les levures qu'il contenait. Il affirmait que le vin vieillirait ainsi sans aigreur. Les industriels appliquèrent, à contrecœur, sa proposition scandaleuse, et constatèrent à leur plus grande joie que l'aigreur avait disparu sans que le chauffage n'ait altéré la saveur du vin.

L'industrie du vin était sauvée. De plus, le procédé consistant à chauffer doucement (*la pasteurisation*) fut ensuite appliqué au lait pour tuer les germes pathogènes.

En dehors des levures, d'autres organismes participent au processus de dégradation. En fait, un processus analogue à la fermentation a lieu dans l'intestin. Le premier à étudier la digestion scientifiquement fut le physicien français René Antoine Ferchault de Réaumur. Son sujet d'expérimentation était l'aigle auquel, en 1752, il fit avaler de petits tubes de métal contenant de la viande ; les tubes étaient prévus pour que la viande ne soit pas broyée, mais ils avaient des ouvertures fermées par des grilles, de sorte que les réactions chimiques ayant lieu dans l'estomac pouvaient s'exercer sur la viande. Réaumur observa que, lorsque l'aigle régurgitait ces tubes, la viande était en partie dissoute et qu'on trouvait dans les tubes un liquide jaunâtre.

En 1777, le médecin écossais Edward Stevens, en extrayant le liquide de l'estomac (*les sucs gastriques*), montra que le processus de dissolution pouvait avoir lieu hors de l'organisme, qu'il n'était donc pas soumis à son influence directe.

Les sucs gastriques contenaient à l'évidence quelque chose qui accélérait la décomposition de la viande. En 1834, le naturaliste allemand Theodor Schwann ajouta du chlorure de mercure au suc gastrique et précipita une poudre blanche. Après élimination du chlorure de mercure de cette poudre blanche, il fit dissoudre ce qui restait et trouva un suc digestif très concentré. Il appela la poudre découverte la *pepsine,* du mot grec signifiant « digérer ».

Pendant ce temps-là, deux chimistes français, Anselme Payen et Jean-François Persoz, avaient découvert dans des extraits de malt une substance qui pouvait provoquer la conversion de l'amidon en sucre plus vite que ne le pouvait l'acide. Ils l'appelèrent *diastase,* du mot grec signifiant « séparer », car ils l'avaient séparée du malt.

Pendant longtemps, les chimistes ont fait une distinction nette entre les ferments vivants tels que la levure, et les ferments non vivants ou *inorganisés,* tels que la pepsine. En 1878, le physiologiste allemand Wilhelm Kühne proposa que ces derniers soient appelés des *enzymes,* du grec signifiant « dans la levure », parce que leur activité était semblable à celle apportée par les substances catalytiques de la levure. Kühne ne se rendait pas compte à quel point le terme « enzyme » allait devenir important, même universel.

En 1897, le chimiste allemand Eduard Buchner broya des levures avec du sable pour casser toutes les cellules et parvint à en extraire un jus. Il trouva que celui-ci pouvait accomplir les mêmes tâches de fermentation que la levure originale. Tout à coup, la distinction entre les ferments à l'intérieur et à l'extérieur des cellules disparaissait ; cette observation allait à l'encontre de la théorie vitaliste de la séparation entre vivant et non-vivant. Le terme « enzyme » est maintenant utilisé pour tous les ferments.

Pour cette découverte, Buchner reçut le prix Nobel de chimie en 1907.

LES CATALYSEURS PROTÉIQUES

On peut dire maintenant qu'une enzyme est simplement un catalyseur organique. Les chimistes commencèrent par isoler des enzymes pour découvrir leur véritable substance. Le problème, c'est que la quantité d'enzyme présente dans les cellules et les sucs naturels est habituellement très faible, et les extraits obtenus sont invariablement des mélanges où

il est difficile de différencier ce qui est dû à l'enzyme de ce qui ne l'est pas.

De nombreux biochimistes soupçonnaient que les enzymes étaient des protéines, parce que les propriétés des enzymes pouvaient être facilement détruites par un léger chauffage, tout comme les protéines. En 1920, le biochimiste allemand Richard Willstätter relata que certaines solutions d'enzyme purifiée, d'où il pensait avoir éliminé toutes les protéines, montraient une activité catalytique très marquée. Il en conclut que les enzymes n'étaient pas des protéines mais des réactifs relativement simples, qui pouvaient en fait utiliser des protéines comme *molécules porteuses.* La plupart des biochimistes suivirent Willstätter qui, lauréat du prix Nobel, bénéficiait d'un grand prestige.

Cependant, à l'université de Cornell, le biochimiste James Batcheller Sumner avança de solides preuves contre cette théorie dès qu'elle fut proposée. A partir de fèves (les graines blanches d'une plante tropicale d'Amérique), Sumner isola des cristaux qui, en solution, avaient les propriétés d'une enzyme appelée l'*uréase,* qui catalyse la dégradation de l'urée en gaz carbonique et ammoniac. Les cristaux de Sumner montraient clairement les propriétés des protéines, et il ne trouva aucun moyen de séparer la protéine de l'activité enzymatique. Tout ce qui dénaturait la protéine détruisait aussi l'enzyme. Tout cela semblait prouver que ce qu'il avait obtenu était une enzyme sous forme pure et cristalline, et que l'enzyme était une protéine.

La grande renommée de Willstätter porta ombrage à la découverte de Sumner pendant un certain temps. Cependant, en mai 1930, le chimiste John Howard Northrop et ses collaborateurs de l'institut Rockefeller confirmèrent l'expérience de Sumner. Ils cristallisèrent un certain nombre d'enzymes, la pepsine incluse, et constatèrent que toutes étaient des protéines. De plus, Northrop montra que ces cristaux étaient des protéines pures ; par ailleurs, ils conservaient leurs propriétés catalytiques même après dissolution et dilution au-delà du point où les tests chimiques ordinaires (les mêmes tests que Willstätter avait utilisés) ne pouvaient plus détecter la présence des protéines.

Les enzymes étaient ainsi considérées comme des *catalyseurs protéiques.* A l'heure actuelle, quelque deux mille enzymes ont été identifiées, et plus de deux cents ont été cristallisées ; ce sont toutes sans exception des protéines. Pour leurs travaux, Sumner et Northrop ont partagé le prix Nobel de chimie en 1946.

L'ACTION ENZYMATIQUE

Les enzymes sont des catalyseurs remarquables à deux égards : l'efficacité et la spécificité. Il existe par exemple une enzyme, la catalase, qui catalyse la dégradation de l'eau oxygénée en eau et oxygène. La dégradation de l'eau oxygénée peut aussi être catalysée par de la limaille de fer ou du bioxyde de manganèse. Cependant, à poids égal, la catalase accélère la vitesse de dégradation beaucoup plus que n'importe quel catalyseur minéral. Chaque molécule de catalase peut causer la dégradation de 44 000 molécules d'eau oxygénée par seconde à 0 °C. Le résultat est qu'une très faible concentration d'enzyme est nécessaire pour accomplir sa fonction.

Pour mettre fin à la vie, il n'est donc besoin que de faibles quantités de substance (*poison*) pour entraver le fonctionnement d'une enzyme clef. Les métaux lourds, quand ils sont administrés sous la forme de chlorure de mercure ou de nitrate de baryum, réagissent avec les groupements thiols, essentiels au fonctionnement de nombreuses enzymes. L'action de ces enzymes s'arrête et l'organisme est empoisonné. Des composés tels que le cyanure de potassium ou le cyanure d'hydrogène se lient par leur groupement cyanure (-CN) à l'atome de fer appartenant à d'autres enzymes essentielles et conduisent à une mort rapide, et on peut l'espérer, sans souffrances ; le cyanure d'hydrogène est en effet le gaz utilisé lors des exécutions dans les chambres à gaz de certains États d'Occident.

L'oxyde de carbone constitue une exception parmi les poisons courants. Il n'agit pas sur une enzyme, mais se lie à la molécule d'hémoglobine (une protéine mais pas une enzyme). Celle-ci transporte normalement l'oxygène des poumons aux cellules, mais elle ne peut le faire quand elle est liée à l'oxyde de carbone. Les animaux qui ne possèdent pas d'hémoglobine ne sont pas empoisonnés par l'oxyde de carbone.

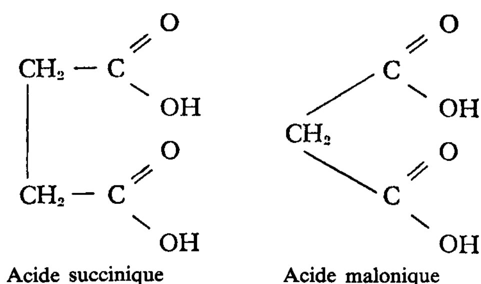
Les enzymes sont extrêmement spécifiques. La catalase en est un bon exemple : elle dégrade l'eau oxygénée et rien d'autre ; tandis que les catalyseurs minéraux, tels que la limaille de fer et le bioxyde de manganèse, peuvent dégrader l'eau oxygénée ainsi que participer à de nombreuses autres réactions.

Qu'est-ce qui explique la spécificité remarquable des enzymes ? La théorie de Lunge et Langmuir, selon laquelle les catalyseurs jouent un rôle d'intermédiaire, donnait une indication. Considérons que l'enzyme a une association temporaire avec le *substrat,* c'est-à-dire la substance dont elle catalyse la réaction. La forme ou la conformation d'une enzyme donnée pourrait alors jouer un rôle très important. Chaque enzyme présente une surface complexe, car elles ont de nombreuses chaînes latérales différentes saillant hors de la chaîne peptidique. Certaines de ces chaînes latérales ont une charge négative, d'autres une charge positive, d'autres sont neutres. Certaines sont volumineuses, d'autres sont petites. On peut imaginer que chaque enzyme a une surface qui ne s'ajuste qu'à un substrat particulier ; chacune est adaptée au substrat comme la clef à la serrure. Elle se liera donc facilement à cette substance, et mal ou pas du tout aux autres. D'où la grande spécificité des enzymes : chacune a une surface faite sur mesure, pour ainsi dire, pour s'associer à un composé déterminé. Dans ce cas, il ne faut pas s'étonner que les protéines soient construites d'unités différentes si nombreuses et soient synthétisées en une si grande variété par le tissu vivant.

Cette théorie de l'activité enzymatique a été proposée en premier par un physiologiste anglais, William Maddock Bayliss, qui travaillait sur une enzyme de la digestion appelée la *trypsine*. En 1913, la théorie fut utilisée par la chimiste allemande Leonor Michaelis et son assistante, Maud Lenora Menten, qui élaborèrent l'équation de Michaelis-Menten ; elles décrivirent la manière dont les enzymes effectuent leurs fonctions et enlevèrent à ces catalyseurs une bonne partie de leur mystère.

Cette vision de l'action enzymatique en tant que serrure et clef fut aussi étayée par la découverte suivante : la présence d'une substance similaire à la structure d'un substrat donné ralentit ou inhibe une enzyme appelée *l'acide succinique déshydrogénase,* qui catalyse l'élimination de deux atomes d'hydrogène de l'acide succinique. Cette réaction n'a pas lieu en présence

d'un composé appelé l'*acide malonique*, qui est tout à fait semblable à l'acide succinique. Les structures des acides succinique et malonique sont :



Acide succinique Acide malonique

La seule différence entre ces deux molécules est que l'acide succinique a un CH_2 supplémentaire à gauche. A cause de sa ressemblance structurale avec l'acide succinique, l'acide malonique peut probablement se lier à la surface de l'enzyme. Une fois que l'acide malonique occupe les lieux à la surface où l'acide succinique se fixe normalement, il reste coincé là, pour ainsi dire, et l'enzyme ne peut plus agir. L'acide malonique « empoisonne » l'enzyme par rapport à son activité normale. Ce type d'activité est appelée une *inhibition compétitive*.

La preuve la plus tangible en faveur de la théorie du complexe enzyme-substrat vient de l'analyse spectrographique. Lorsqu'une enzyme s'associe à son substrat, on devrait probablement observer une modification du spectre d'absorption : l'absorption de la lumière du complexe devrait être différente de celle de l'enzyme ou du substrat seul. En 1936, les biochimistes anglais David Keilin et Thaddeus Mann détectèrent un changement de couleur dans une solution d'enzyme, la peroxydase, après addition de son substrat, l'eau oxygénée. Le biophysicien américain Britton Chance fit une analyse spectrale, et trouva qu'il y avait deux changements progressifs dans le mode d'absorption, l'un suivant l'autre. Il attribua le premier changement à la formation, à une certaine vitesse, du complexe enzyme-substrat, et le second au déclin de cette association quand la réaction s'achève. En 1964, le biochimiste japonais Kunio Yagi annonça l'isolement d'un complexe enzyme-substrat fait de l'union lâche entre l'enzyme D-amino-oxydase et son substrat, l'alanine.

Maintenant, on peut se poser la question suivante : la molécule entière de l'enzyme est-elle nécessaire à la catalyse, ou une partie seulement est-elle suffisante ? C'est une question importante d'un point de vue pratique et théorique. Les enzymes sont largement utilisées de nos jours ; on les a beaucoup employées pour la fabrication de médicaments, de l'acide citrique et de nombreux autres produits chimiques. Si la molécule d'enzyme n'est pas nécessaire dans son entier et que de petits fragments de cette molécule suffisent à faire le travail, on pourrait synthétiser la partie active ; ainsi, le processus n'aurait pas à dépendre de l'utilisation de cellules vivantes, telles que les levures, les champignons ou les bactéries.

On a fait des progrès prometteurs dans ce sens. Par exemple, Northrop a découvert que si des groupements acétyles (CH_3CO-) sont fixés sur les chaînes latérales de l'acide aminé tyrosine de la molécule de pepsine, celle-ci perd une partie de son activité. En revanche, il n'y a pas de perte d'activité quand des groupements acétyles sont fixés à la chaîne latérale de la lysine de la pepsine. La tyrosine doit donc participer à l'activité de

la pepsine, tandis que la lysine, évidemment, n'y participe pas. Cela montrait pour la première fois que certaines portions de la molécule d'enzyme ne sont pas essentielles à son activité.

Le *site actif* d'une autre enzyme de la digestion a été décrit récemment avec plus de précision. Cette enzyme est la chymotrypsine. Le pancréas la sécrète d'abord sous une forme inactive appelée le *chymotrypsinogène.* Cette forme inactive est convertie dans la forme active par la coupure d'une seule liaison peptidique (ce qui est accompli par une autre enzyme de la digestion, la trypsine) ; il semble donc que le démasquage d'un seul acide aminé convertisse la chymotrypsine en une enzyme active. Il s'avère que la fixation, sur la molécule de chymotrypsine, d'une molécule connue sous le sigle DFP (*diisopropyl fluorophosphate*) bloque son activité enzymatique. Il est probable que le DFP se lie à un acide aminé essentiel. Grâce à son marquage par le DFP, cet acide aminé a été identifié ; c'est une sérine. En fait, on a aussi trouvé que le DFP se lie à la sérine d'autres enzymes digestives. Dans chaque cas, la sérine occupe la même position dans une séquence de quatre acides aminés : glycine-acide aspartique-sérine-glycine.

En réalité, un peptide constitué seulement de ces quatre acides aminés ne possède pas d'activité catalytique. Le reste de la molécule joue aussi un rôle certain. On peut penser que la séquence de quatre acides aminés du centre actif est comme la lame d'un couteau qui, sans manche, ne sert à rien.

Le site actif, ou la lame, n'existe pas forcément non plus en un seul morceau sur la chaîne d'acides aminés. Considérons une enzyme : la ribonucléase. Maintenant que nous connaissons l'ordre exact de ses 124 acides aminés, il est possible de concevoir des méthodes pour altérer délibérément tel ou tel acide aminé de la chaîne et de noter l'effet du changement sur l'activité enzymatique. On a découvert ainsi que trois acides aminés sont plus particulièrement nécessaires à son activité, bien que très éloignés les uns des autres. Ce sont : l'histidine en position 12, la lysine en position 41, et une autre histidine en position 119.

Cette séparation n'existe, bien entendu, que sur la chaîne vue comme une longue corde. Dans la molécule en état de fonctionnement, la chaîne est enroulée en une configuration tridimensionnelle spécifique, maintenue en place par quatre molécules de cystine placées en travers des boucles. Dans une telle molécule, les trois acides aminés nécessaires sont rassemblés en une unité compacte.

La composition du centre actif a été mieux définie encore dans le cas du lysozyme, une enzyme trouvée en de nombreux endroits, dont les larmes et le mucus nasal. Il détruit les bactéries en catalysant la dégradation de certaines substances qui constituent la paroi bactérienne. C'est comme s'il fissurait la paroi bactérienne, permettant alors à son contenu de s'échapper.

Le lysozyme est la première enzyme dont la structure a été complètement élucidée (en 1965) en trois dimensions. Cela fait, on pouvait montrer que la molécule de la paroi bactérienne, qui est sujette à l'action du lysozyme, s'ajuste correctement le long d'une faille organisée dans la structure de l'enzyme. La liaison clef se fait entre l'atome d'oxygène de la chaîne latérale de l'acide glutamique (position 35) et un autre atome d'oxygène dans la chaîne latérale de l'acide aspartique en

position 52. Les deux positions sont rapprochées par le repliement de la chaîne d'acides aminés, laissant un espace juste assez grand pour que la molécule attaquée se loge entre elles. Les réactions chimiques nécessaires pour rompre la liaison peuvent facilement avoir lieu dans ces conditions ; le lysozyme est particulièrement bien organisé pour accomplir sa tâche.

Il arrive aussi quelquefois que la lame de l'enzyme ne soit pas formée d'un groupe d'acides aminés, mais d'une association d'atomes de nature entièrement différente. Nous en parlerons plus loin.

Si nous ne pouvons pas toucher à la lame, pouvons-nous modifier le manche sans altérer l'utilité de l'outil ? Le manche a ses raisons d'être, bien sûr. Il semblerait que l'enzyme, dans son état naturel, soit flexible et puisse prendre plusieurs formes différentes sans trop de contrainte. Quand le substrat se fixe au site actif, l'enzyme s'ajuste à la forme du substrat, grâce au « jeu » de la portion inactive de la molécule, si bien que l'assemblage est bloqué et que l'action catalytique est très efficace. Les substrats de forme légèrement différente peuvent ne pas tirer aussi bien leur avantage du « jeu » et ne seront pas modifiés ; ils peuvent même inhiber l'enzyme.

Cependant, l'enzyme peut être simplifiée, peut-être au prix d'une certaine perte d'efficacité. L'existence d'une protéine sous différentes formes, comme l'insuline par exemple, nous encourage à penser qu'une simplification est possible. L'insuline est une hormone, pas une enzyme, mais ses fonctions sont très spécifiques. A une certaine position sur la chaîne G, on trouve une séquence de trois acides aminés qui est différente selon l'animal : chez les bovins, elle est alanine-sérine-valine ; chez le porc, thréonine-sérine-isoleucine ; chez le mouton, alanine-glycine-valine ; chez le cheval, thréonine-glycine-isoleucine, etc. Pourtant, chacune de ces insulines peut être substituée à une autre, tout en accomplissant la même fonction.

Qui plus est, une molécule de protéine peut parfois être raccourcie considérablement sans aucun effet important sur son activité (comme le manche d'un couteau ou d'une hache peut être raccourci sans trop de perte d'efficacité). Un cas d'espèce est l'hormone appelée ACTH (*hormone adrénocorticotrope*). C'est une chaîne peptidique constituée de trente acides aminés, dont l'ordre a maintenant été complètement déterminé. On a pu éliminer jusqu'à quinze acides aminés de l'extrémité C terminale, sans détruire l'activité de l'hormone. Mais l'élimination de un ou deux acides aminés de l'extrémité N terminale (la lame, pour ainsi dire) détruit l'activité d'un seul coup.

On a répété ces expériences sur une enzyme appelée la *papaïne*, provenant du fruit et de la sève du papayer. Son action enzymatique est semblable à celle de la pepsine. L'élimination de 180 acides aminés de la partie N terminale de la molécule de pepsine ne réduit pas son activité de façon notable.

Il est donc au moins concevable de simplifier encore les enzymes jusqu'à pouvoir les synthétiser massivement. Les enzymes synthétiques, sous la forme de composés relativement simples, pourraient alors être fabriquées sur une grande échelle, en vue d'applications variées. Ce serait une sorte de miniaturisation chimique.

Le métabolisme

Un organisme tel que le corps humain est une usine chimique aux multiples facettes. Il respire de l'oxygène et boit de l'eau. Il prend dans sa nourriture des sucres (glucides), des graisses, des protéines, des minéraux et autres matériaux bruts. Il élimine les produits indigestes, en plus des bactéries et des produits de putréfaction qu'elles engendrent. Il excrète aussi du gaz carbonique par les poumons, rend de l'eau par la respiration et la transpiration, il excrète de l'urine qui transporte de nombreux composés en solution, le principal étant l'urée. Ces réactions chimiques déterminent le *métabolisme* du corps.

En examinant les matériaux bruts qui entrent dans le corps et les déchets qui en sortent, on peut glaner des informations sur ce qui s'y passe. Par exemple, étant donné que les protéines fournissent la majeure partie de l'azote qui entre dans le corps, nous savons que l'urée (NH_2CONH_2) doit être un produit du métabolisme des protéines. Mais entre les protéines et l'urée il y a une route longue, tortueuse et compliquée. Chaque enzyme du corps ne catalyse qu'une petite réaction spécifique, ne modifiant le plus souvent que deux ou trois atomes. Chaque conversion majeure dans le corps implique une multitude d'étapes et de nombreuses enzymes. Même un organisme apparemment aussi simple que la minuscule bactérie doit utiliser plusieurs milliers d'enzymes et de réactions différentes.

Tout cela peut sembler inutilement complexe, mais c'est la véritable essence de la vie. Le vaste complexe de réactions des tissus peut être contrôlé avec une très grande précision en augmentant ou en diminuant la production des enzymes appropriées. Les enzymes gouvernent la chimie du corps, comme le mouvement compliqué des doigts sur les cordes gouverne le jeu du violon ; et sans cette complexité, le corps ne pourrait pas accomplir ses fonctions multiples.

Suivre le cours des multiples réactions qui constituent le métabolisme du corps, c'est suivre le cours de la vie. Tenter de le suivre en détail, pour comprendre l'interdépendance des innombrables réactions qui ont lieu toutes en même temps, peut sembler une entreprise terrifiante, voire désespérée.

LA CONVERSION DU SUCRE EN ALCOOL ÉTHYLIQUE

Pour comprendre le métabolisme, les chimistes ont commencé par étudier la manière dont les cellules de levure convertissent le sucre en alcool éthylique. En 1905, deux chimistes britanniques, Arthur Harden et William John Young, ont émis l'hypothèse que ce processus passait par la formation de sucres portant des groupements phosphates. Harden et Young ont été les premiers à remarquer que le phosphore joue un rôle important dans le métabolisme (et le phosphore a pris depuis lors un rôle de premier plan). Harden et Young ont même trouvé dans le tissu vivant un ester de sucre et de phosphate constitué de fructose, sur lequel sont liés deux groupements phosphates (PO_3H_2). Ce *fructose diphosphate* (encore connu sous le nom d'ester de Harden-Young) a été le premier *intermédiaire métabolique* à être identifié d'une façon définitive – le premier composé, donc, reconnu comme étant formé au cours du processus menant des

composés ingérés par l'organisme aux composés éliminés par celui-ci. Pour ces travaux et d'autres concernant la conversion du sucre en alcool par la levure (voir chapitre 15), Harden partagea le prix Nobel de chimie en 1929.

Ce qui ne concernait au début que la levure prit une importance beaucoup plus grande quand le chimiste allemand Otto Fritz Meyerhof démontra, en 1918, que les cellules animales, telles que celles du muscle, dégradent le sucre de la même façon que le fait la levure ; la différence principale étant que, dans les cellules animales, la dégradation ne va pas aussi loin dans cette voie métabolique. Au lieu de convertir la molécule de glucose à six carbones en une molécule d'alcool éthylique (CH_3CH_2OH) à deux carbones, elles le dégradent seulement en acide lactique qui a trois carbones ($CH_3CHOHCOOH$).

Les travaux de Meyerhof venaient confirmer pour la première fois le principe général, qui a été depuis largement accepté, selon lequel le métabolisme suit les mêmes voies, à part quelques différences mineures, chez toutes les créatures, des plus simples aux plus complexes. Les études de Meyerhof sur l'acide lactique dans le muscle lui valurent de partager le prix Nobel de physiologie et de médecine en 1922 avec le physiologiste anglais Archibald Vivian Hill. Ce dernier avait abordé le muscle du point de vue de la production de chaleur, et était arrivé à des conclusions tout à fait similaires à celles que Meyerhof avaient obtenues par la voie chimique.

Le détail des étapes impliquées dans le passage du sucre à l'acide lactique fut étudié de 1937 à 1941 par Carl Ferdinand Cori et sa femme Gerty Theresa Cori, qui travaillaient à l'université Washington à Saint Louis. Ils ont utilisé des extraits tissulaires et purifié les enzymes qui provoquent des modifications de divers esters phosphoriques des sucres, puis ils ont ordonné les étapes à la manière d'un puzzle. Le schéma des modifications, étape par étape, qu'ils ont présenté, tient encore debout de nos jours, à quelques détails près, et le prix Nobel de physiologie et de médecine leur fut attribué en 1947.

Lors du passage du sucre à l'acide lactique, une certaine quantité d'énergie est produite et utilisée par les cellules. Les cellules de levure se servent de cette énergie lors de la fermentation du sucre, comme le fait la cellule musculaire, quand c'est nécessaire. Il est important de se souvenir que l'énergie est obtenue sans utiliser l'oxygène de l'air. Donc, un muscle est capable de travailler, même quand il doit dépenser une énergie supérieure à celle qu'il remplacerait par des réactions impliquant l'oxygène qui lui est apporté à une vitesse relativement lente par le sang. Comme l'acide lactique s'accumule, le muscle se fatigue, et en fin de compte il doit se reposer jusqu'à ce que l'oxygène lui permette de dégrader l'acide lactique.

L'ÉNERGIE MÉTABOLIQUE

La question qui se pose maintenant est la suivante : sous quelle forme l'énergie libérée par la dégradation du sucre en acide lactique est-elle fournie aux cellules, et comment celles-ci l'utilisent-elles ? Le chimiste américain né en Allemagne, Fritz Albert Lipmann, a trouvé la réponse au cours de ses recherches qui ont commencé en 1941. Il a montré que

certains composés phosphorylés formés au cours du métabolisme des glucides emmagasinent une quantité d'énergie inhabituelle dans la liaison qui relie le groupement phosphate au reste de la molécule. Cette *liaison phosphate à haute énergie* est transmise à des transporteurs d'énergie présents dans toutes les cellules. Le transporteur le plus connu est l'*adénosine triphosphate* (ATP). La molécule d'ATP et certains composés similaires représentent la « petite monnaie » de l'énergie de l'organisme. Elle conserve l'énergie dans des unités de taille pratique et aisément utilisables. Quand la liaison phosphate est hydrolysée, l'énergie libérée est convertie soit en énergie chimique permettant la formation de protéines à partir d'acides aminés, soit en énergie électrique pour la transmission de l'influx nerveux, soit en énergie cinétique par l'intermédiaire de la contraction musculaire, et ainsi de suite. Bien que la quantité d'ATP présente dans le corps soit faible à un temps donné, il y en a toujours assez (tant que la vie continue), car l'ATP est reformé aussi vite qu'il est utilisé.

Pour cette découverte capitale, Lipmann partagea le prix Nobel de physiologie et de médecine en 1953.

L'organisme des mammifères ne peut pas convertir l'acide lactique en alcool éthylique (comme le fait la levure) ; utilisant une autre voie métabolique, l'organisme évite l'alcool éthylique, et dégrade l'acide lactique directement en gaz carbonique (CO_2) et eau. Pour ce faire, il consomme de l'oxygène et produit beaucoup plus d'énergie qu'en convertissant le glucose en acide lactique en l'absence d'oxygène.

On dispose d'un moyen facile pour suivre un processus métabolique qui comprend une consommation d'oxygène : il suffit de déterminer quels sont les composés intermédiaires créés en cours de route. Supposons qu'à une étape donnée d'une séquence de réactions, une certaine substance (l'acide succinique, par exemple) soit un substrat intermédiaire. On peut mélanger cet acide à du tissu vivant (ou souvent à une seule enzyme) et mesurer la vitesse de la consommation de l'oxygène de ce mélange. S'il montre une consommation élevée d'oxygène, on peut penser avec certitude que la substance participe à la réaction.

Le biochimiste allemand Otto Heinrich Warburg a conçu l'instrument qui a servi à mesurer la vitesse de consommation de l'oxygène. C'est le *manomètre de Warburg*, qui consiste en un petit flacon dans lequel le substrat et le tissu (ou bien l'enzyme) sont mélangés. Ce flacon est relié à l'une des extrémités d'un tube fin en U, dont l'autre extrémité est ouverte. Un liquide coloré remplit la partie basse du U. Comme le mélange enzyme-substrat absorbe l'oxygène du flacon, un léger vide se produit, et le liquide coloré du tube en U monte dans la partie du tube connectée au flacon. La vitesse à laquelle le liquide monte est utilisée pour calculer la vitesse de consommation de l'oxygène (figure 12.4).

Les expériences de Warburg sur la consommation d'oxygène par les tissus lui valurent le prix Nobel de physiologie et de médecine en 1931.

Warburg et un autre biochimiste allemand, Heinrich Wieland, identifièrent les réactions qui fournissent de l'énergie pendant la dégradation de l'acide lactique. Au cours d'une série de réactions, des paires d'atomes d'hydrogène sont éliminées de substances intermédiaires au moyen d'enzymes appelées *déshydrogénases*. Ces atomes d'hydrogène se combinent avec l'oxygène, grâce à la catalyse enzymatique des *cytochromes*. A la fin des années vingt, Warburg et Wieland ont discuté avec acharnement pour savoir laquelle des deux réactions était la plus importante, Warburg

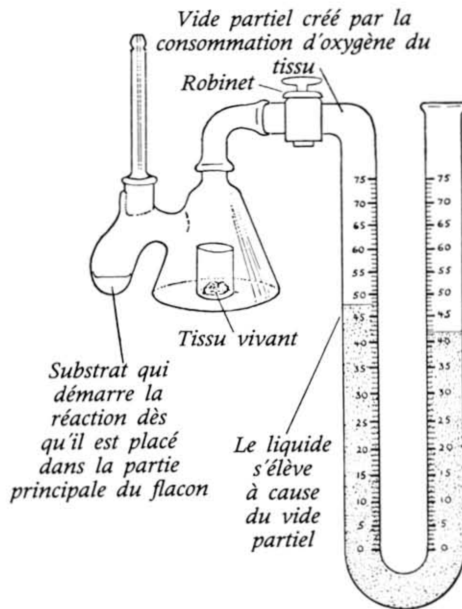


Figure 12.4. Le manomètre de Warburg.

soutenant que c'était la capture d'oxygène, et Wieland que c'était l'élimination de l'hydrogène. Finalement, David Keilin démontra que les deux étapes sont essentielles.

Le biochimiste allemand Hans Adolf Krebs a établi la séquence des réactions et des produits intermédiaires, lors du passage de l'acide lactique au gaz carbonique et à l'eau. Il s'agit du *cycle de Krebs,* ou *cycle tricarboxylique,* l'acide citrique, un triacide, étant l'un des composés principaux de cette voie métabolique. Pour cette réalisation, achevée en 1940, Krebs partagea avec Lipmann le prix Nobel de physiologie et de médecine en 1953.

Le cycle de Krebs produit la part du lion de l'énergie générée par les organismes qui utilisent l'oxygène moléculaire pour leur respiration (c'est-à-dire tous les organismes sauf quelques bactéries anaérobiques, qui tirent leur énergie de réactions chimiques indépendantes de l'oxygène). En différents endroits du cycle de Krebs, un composé perd deux atomes d'hydrogène, qui finissent – au cours d'un processus complexe – par se lier à l'oxygène pour former de l'eau. Les deux atomes d'hydrogène passent d'une sorte de cytochrome à l'autre, jusqu'à la dernière, *la cytochrome oxydase,* qui les transfère sur l'oxygène moléculaire. Tout au long de la chaîne de cytochromes, des molécules d'ATP sont formées, et l'organisme est alimenté par cette « petite monnaie » d'énergie. A chaque tour du cycle de Krebs, dix-huit molécules d'ATP sont formées. Le processus dans son entier, impliquant l'oxygène et la fixation de groupements phosphates pour former l'ATP, est appelé *l'oxydation phosphorylante.* C'est une série de réactions capitale pour les tissus vivants. Toute interférence sérieuse bloquant ce processus (comme lorsqu'on absorbe du cyanure de potassium) provoque la mort en quelques minutes.

Des granules minuscules, qui sont dans le cytoplasme, contiennent toutes les substances et les enzymes participant aux oxydations phosphorylantes. Elles ont été mises en évidence en 1898 par le biologiste allemand C. Benda, qui ne comprenait pas leur importance à l'époque. Il les appela *mitochondries,* car il pensait, à tort, que c'étaient des « fils de cartilage », et le nom leur est resté depuis.

Les mitochondries ont généralement la forme d'un ballon de rugby, sont longues de 2,5 microns et ont 1 micron d'épaisseur. Les cellules en contiennent de plusieurs centaines à un millier, en moyenne. Les très grandes cellules peuvent en contenir deux cent mille, tandis que les bactéries anaérobiques n'en contiennent pas du tout. Après la Deuxième Guerre mondiale, des observations de microscopie électronique ont montré que, malgré leur taille minuscule, les mitochondries ont une structure complexe qui leur est propre. La mitochondrie a une double membrane, la membrane extérieure étant lisse et la membrane interne plissée, de telle façon qu'elle offre une grande surface. Tout au long de la membrane interne, on trouve des *particules élémentaires,* minuscules, qui semblent être le siège des oxydations phosphorylantes.

LE MÉTABOLISME DES GRAISSES

A la même époque, les biochimistes ont également progressé dans leur compréhension du métabolisme des graisses. On savait que les molécules de graisse sont des chaînes de carbones, qu'elles peuvent être hydrolysées en *acides gras* (les plus courants ont seize ou dix-huit atomes de carbone), et que ces molécules sont dégradées par deux carbones à la fois. En 1947, Fritz Lipmann découvrit un composé complexe, le *coenzyme A* (A pour acétylation), qui joue un rôle dans *l'acétylation,* c'est-à-dire le transfert d'un chaînon de deux carbones d'un composé à un autre. Trois ans plus tard, le biochimiste allemand Feodor Lynen trouva que le coenzyme A était impliqué dans la dégradation des acides gras. La fixation d'un acide gras sur le coenzyme A est suivie par une série de quatre étapes se terminant par l'élimination des deux carbones situés à l'extrémité de la chaîne d'acides gras liée au coenzyme A. Ce qui reste de l'acide gras se lie alors à une autre molécule de coenzyme A, deux atomes de carbone de plus sont éliminés, etc. C'est le *cycle d'oxydation des acides gras.* Cela valut à Lynen de partager le prix Nobel de physiologie et de médecine en 1964.

La dégradation des protéines est souvent plus compliquée que celle des glucides ou des graisses, car elle implique quelque vingt acides aminés différents. Dans certains cas néanmoins, les choses sont relativement simples : la modification mineure d'un acide aminé le convertit en un composé qui peut entrer dans le cycle tricarboxylique (comme le font les chaînons dicarbonés formés à partir des acides gras). Cependant, les acides aminés sont décomposés selon des voies complexes.

Nous pouvons maintenant retourner à la conversion des protéines en urée – point annoncé dans la section des enzymes. Cette conversion est relativement simple.

La chaîne latérale de l'arginine contient un groupement qui est essentiellement celui de l'urée. Ce groupement peut être détaché par une enzyme, l'*arginase,* laissant une sorte d'acide aminé tronqué, l'*ornithine.* En 1932, en étudiant la formation de l'urée par des tissus de rat, Krebs et

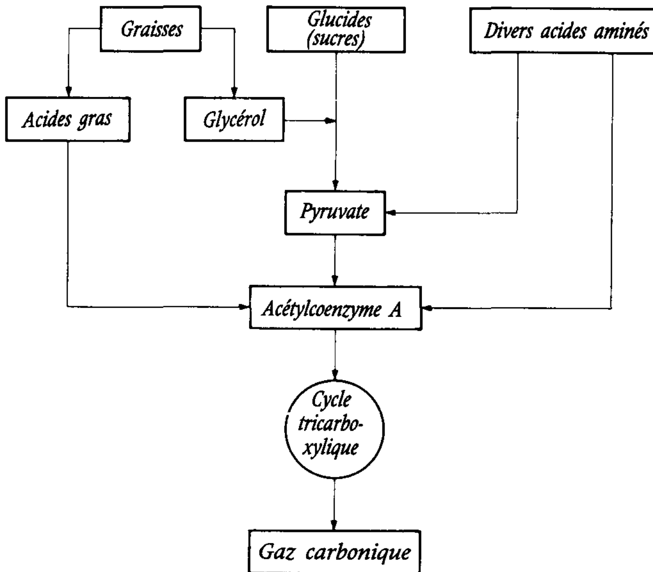
un collaborateur, K. Henseleit, découvrirent qu'après addition d'arginine aux tissus, un flux d'urée était produit, beaucoup plus important que celui qui résulterait seulement de la coupure de toutes les molécules d'arginine ajoutées. Krebs et Henseleit pensèrent que les molécules d'arginine fonctionnent comme un agent qui produit de l'urée sans arrêt. En d'autres termes, une fois l'arginine débarrassée de l'urée grâce à l'arginase, l'ornithine libérée se charge d'un nouveau groupement aminé provenant d'autres acides aminés (en plus du gaz carbonique provenant de l'organisme) et reforme ainsi de l'arginine. La molécule d'arginine est donc régulièrement dégradée, reformée, dégradée, etc., donnant chaque fois une molécule d'urée. Ce processus est appelé *cycle de l'urée, cycle de l'ornithine,* ou encore *cycle de Krebs-Henseleit.*

Après l'élimination de l'azote au moyen de l'arginine, le squelette carboné de l'acide aminé qui reste est dégradé en gaz carbonique et en eau par diverses voies produisant de l'énergie. (Pour se faire une idée d'ensemble du métabolisme des glucides, des graisses et des protéines, se reporter à la figure 12.5).

Les traceurs

Malgré toutes les recherches effectuées sur le métabolisme, les biochimistes n'avaient encore qu'une vue extérieure des choses. Ils pouvaient élucider les voies métaboliques dans leurs grandes lignes, mais

Figure 12.5. Schéma général du métabolisme des sucres, des graisses et des protéines.



pour découvrir ce qui se passait réellement dans l'animal vivant, ils avaient besoin d'un moyen de suivre en détail le déroulement des événements lors des différentes étapes du métabolisme, c'est-à-dire de suivre la destinée des molécules là où elles se trouvent. En fait, les techniques nécessaires avaient été découvertes au début du siècle, mais les chimistes ont mis du temps à en tirer parti.

Le biochimiste allemand Franz Knoop fut le premier à entreprendre ces recherches. En 1904, il conçut l'idée de nourrir des chiens avec des acides gras marqués, pour voir ce qui leur arrivait. Il les marqua en fixant un atome de benzène à l'une des extrémités de la chaîne ; il utilisa le noyau de benzène, car les mammifères ne possèdent pas d'enzymes capables de le dégrader. Knoop espérait que ce qui était lié au benzène lors de son élimination par les urines lui procurerait des informations sur la manière dont les graisses sont dégradées par l'organisme. Il avait raison. On retrouvait toujours le benzène lié à un chaînon de deux carbones. Il en a déduit que l'organisme devait détacher les carbones deux par deux des acides gras. Comme nous l'avons vu, plus de quarante ans plus tard, les travaux sur le coenzyme A confirmèrent ses déductions.

Les chaînes carbonées des acides gras ordinaires contiennent toutes un nombre pair d'atomes de carbone. Que se passe-t-il si la chaîne en a un nombre impair ? Dans ce cas-là, si les atomes de carbone sont détachés par deux, on devrait retrouver un seul atome de carbone lié au benzène. Knoop fit ingérer ce type de graisse aux chiens et en effet, il trouva ce composé.

Knoop avait employé le premier *traceur* de la biochimie. En 1913, le chimiste hongrois Georg von Hevesy et son collègue, le chimiste allemand Friedrich Adolf Paneth, inventèrent un autre moyen de marquer les molécules : les isotopes radioactifs. Ils commencèrent par le plomb, et déterminèrent quelle quantité de plomb une plante pouvait fixer sous la forme de sels de plomb. La quantité était sûrement trop petite pour être mesurable à l'aide des méthodes chimiques disponibles, mais en utilisant du plomb radioactif, on pouvait aisément le doser par sa radioactivité. Hevesy et Paneth nourrirent des plantes avec des solutions de sels de plomb marqué radioactivement ; à intervalles réguliers, ils brûlaient les plantes et mesuraient la radioactivité de leurs cendres. Ils déterminèrent ainsi la vitesse d'absorption du plomb par les cellules des plantes.

Mais le benzène et le plomb sont des substances bien peu « physiologiques » pour être utilisées comme marqueurs. Elles peuvent facilement modifier la chimie normale des cellules vivantes. Il est préférable d'utiliser des atomes marqueurs qui participent vraiment au métabolisme normal de l'organisme, tels que les atomes d'oxygène, d'azote, de carbone, d'hydrogène et de phosphore.

Après que la radioactivité artificielle fut mise en évidence par les Joliot-Curie en 1934, Hevesy continua aussitôt dans cette direction et utilisa des composés phosphorylés contenant du phosphore radioactif. Il mesura ainsi l'incorporation du phosphate dans les plantes. Malheureusement, les radio-isotopes des éléments principaux des tissus vivants – notamment l'azote et l'oxygène – ne sont pas utilisables, car leur temps de demi-vie est de quelques minutes au mieux. Mais les éléments les plus importants ont des isotopes *stables* qui sont utilisables comme marqueurs. Ces isotopes sont le carbone 13, l'azote 15, l'oxygène 18 et l'hydrogène 2. On ne les trouve normalement qu'en très faible quantité (1 % ou moins environ).

Donc, si on « enrichit » l'hydrogène naturel en hydrogène 2, par exemple, on peut s'en servir en tant que marqueur distinctif d'une molécule fournie à l'organisme. La présence d'hydrogène lourd dans un composé est détectée à l'aide d'un spectrographe de masse, qui sépare ce composé des autres en raison de sa masse supérieure. La destinée de l'hydrogène marqué peut donc être suivie à travers l'organisme.

L'hydrogène a en fait servi de premier traceur physiologique, parce qu'il devint disponible pour cet usage lorsque Harold Urey isola l'hydrogène 2 (deutérium), en 1931. L'une des premières choses mises en évidence par l'utilisation du deutérium en tant que traceur fut que les atomes d'hydrogène sont fixés moins solidement à leurs composés, dans l'organisme, qu'on le supposait jusqu'alors. Ils font la navette d'un composé à l'autre. Ils vont des atomes d'oxygène des molécules de sucre à ceux des molécules d'eau, etc. Étant donné qu'on ne peut distinguer un atome d'hydrogène ordinaire d'un autre, ce va-et-vient n'a pu être détecté que lorsque les atomes de deutérium l'ont révélé. Cette découverte implique que les atomes d'hydrogène se déplacent dans tout l'organisme et que, donc, les atomes de deutérium, liés à l'oxygène, se répartissent dans tout l'organisme – que les composés subissent des modifications chimiques globales ou non. Par conséquent, le chercheur doit s'assurer qu'un atome de deutérium trouvé dans un composé provient bien d'une réaction enzymatique précise, et non pas simplement du processus de va-et-vient ou d'échange. Heureusement, les atomes d'hydrogène fixés aux atomes de carbone ne sont pas échangeables, de sorte que le deutérium trouvé sur les chaînes carbonées a une signification métabolique.

Les habitudes vagabondes des atomes ont été soulignées davantage encore quand, en 1937, le biochimiste américain d'origine allemande Rudolf Schoenheimer et ses associés utilisèrent l'azote 15. Ils nourrirent des rats avec des acides aminés marqués à l'azote 15, les tuèrent à des moments déterminés, et analysèrent les tissus pour déterminer les composés dans lesquels on retrouvait de l'azote 15. Là aussi, on remarqua que les réactions d'échange étaient importantes. Après qu'un acide aminé marqué a pénétré l'organisme, on retrouve rapidement l'azote 15 sur presque tous les acides aminés. En 1942, Schoenheimer publia un livre, *The Dynamic State of Body Constituents* (l'état dynamique des constituants de l'organisme). Ce titre décrit le regard neuf sur la biochimie que les traceurs isotopiques ont apporté. Les atomes sont dans un état perpétuel d'agitation en dehors de toute modification chimique réelle.

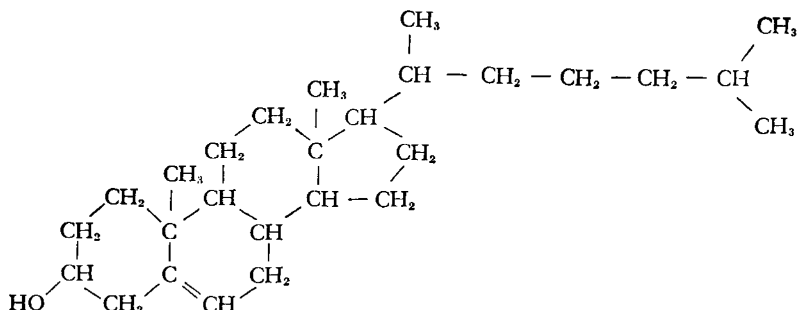
Petit à petit, l'utilisation des traceurs permit de comprendre les détails des voies métaboliques. Le modèle général de la dégradation des sucres, du cycle tricarboxylique et de l'urée fut ainsi confirmé. Il en résulta aussi l'addition de nouveaux intermédiaires, l'établissement de voies alternatives, etc.

Grâce aux réacteurs nucléaires, après la Deuxième Guerre mondiale, plus d'une centaine d'isotopes radioactifs furent rendus disponibles et les travaux au moyen de traceurs devinrent très nombreux. Des composés ordinaires pouvaient être bombardés par des neutrons, et en ressortir chargés d'isotopes radioactifs. Presque tous les laboratoires de biochimie aux États-Unis entreprirent des recherches au moyen de traceurs radioactifs (je pourrais même dire dans le monde entier, puisque les États-Unis ont rapidement mis les isotopes à la disposition des autres États pour usage scientifique).

Au nombre des traceurs stables et radioactifs, on trouve maintenant l'hydrogène (tritium), le phosphore (phosphore 32), le soufre (soufre 35), le potassium (potassium 42), le sodium, l'iode, le fer, le cuivre, et le plus important de tous, le carbone (carbone 14). Le carbone 14 a été découvert en 1940 par les chimistes américains Martin David Kamen et Samuel Ruben qui, à leur grande surprise, trouvèrent qu'il avait une demi-vie de 5 000 ans, ce qui est particulièrement long pour le radio-isotope d'un élément léger.

LE CHOLESTÉROL

Le carbone 14 a permis de résoudre des problèmes qui ont mis au défi les chimistes pendant des années, et auxquels ils semblaient ne pas pouvoir trouver de solution. L'une des énigmes à laquelle il apporta une amorce de réponse fut la formation d'une substance connue sous le nom de *cholestérol.* La formule du cholestérol, définie après de nombreuses années de recherches laborieuses par des hommes tels que Wieland (qui reçut le prix Nobel de chimie en 1927 pour ses travaux sur des composés apparentés au cholestérol) est la suivante :



On ne connaît toujours pas très bien la fonction du cholestérol dans l'organisme, mais elle est importante. Le cholestérol existe en grande quantité dans la gaine graisseuse autour des nerfs, dans les glandes surrénales, et associé à certaines protéines. En excès, il peut être la cause de calculs biliaires et d'artériosclérose. Le cholestérol est le prototype de toute la famille des *stéroïdes,* le noyau des stéroïdes étant constitué de l'association des quatre cycles que l'on voit dans la formule. Les stéroïdes constituent un groupe de substances solides et grasses, qui comprennent les hormones sexuelles et les hormones surrénales. Elles dérivent toutes, sans aucun doute possible, du cholestérol. Comment le cholestérol est-il synthétisé dans l'organisme ?

Les biochimistes n'en avaient pas la moindre idée jusqu'à ce que les traceurs arrivent à leur secours. Les premiers à s'attaquer à cette question ont été Rudolf Schoenheimer et son collaborateur David Rittenberg. Ils firent boire de l'eau lourde à des rats et trouvèrent du deutérium dans les molécules de cholestérol. Cet effet n'était pas significatif en lui-même, parce que le deutérium pouvait provenir des réactions d'échange. Mais, en 1942 (après le suicide de Schoenheimer), Rittenberg et un autre collaborateur, le germano-américain Konrad Emil Bloch, aboutirent à une indication plus précise. Ils donnèrent à manger à des rats des ions acétates

(un groupement simple de deux carbones, CH_3COO_-) marqués au deutérium sur le groupement CH_3. Ils retrouvèrent de nouveau le deutérium dans les molécules de cholestérol, mais cette fois-ci il ne pouvait y être parvenu par simple échange : il avait inévitablement été incorporé dans la molécule, associé au groupement CH_3.

Les groupements de deux carbones (dont l'ion acétate n'est qu'une version) semblent être une plaque tournante du métabolisme. De tels groupements peuvent donc parfaitement servir de réservoirs de matériaux de construction pour le cholestérol. Mais comment constituent-ils cette molécule ?

En 1950, quand le carbone 14 fut rendu disponible, Bloch répéta l'expérience, repérant différemment cette fois les deux carbones de l'ion acétate. Il marqua le carbone du groupement CH_3 avec le carbone 13, qui est un traceur stable, et le carbone du COO_- avec le carbone 14, qui est radioactif. Après avoir fait ingérer le composé par des rats, il analysa leur cholestérol pour voir où les deux carbones marqués se retrouvaient dans la molécule. Cette analyse réclamait une chimie sophistiquée, et Bloch, ainsi que de nombreux autres expérimentateurs, y travaillèrent pendant des années, identifiant chacun des carbones du cholestérol l'un après l'autre. Le résultat en a été que les groupements acétates formaient probablement d'abord une molécule appelée *squalène*, un composé de trente carbones relativement rare dans l'organisme, auquel personne n'avait sérieusement prêté attention jusque-là. Il apparaît maintenant que le squalène est une étape dans la voie de la synthèse du cholestérol. Les biochimistes commencèrent à l'étudier avec un intérêt intense. Bloch partagea le prix Nobel de physiologie et de médecine avec Lynen en 1964.

LE CYCLE PORPHYRIQUE DE L'HÈME

Les biochimistes ont cherché à comprendre la construction du cycle porphyrique de l'hème, structure centrale de la molécule d'hémoglobine et de bien d'autres, comme ils l'avaient fait pour la synthèse du cholestérol. David Shemin, de l'université Columbia à New York, nourrit des canards avec de la glycine marquée de différentes façons. La glycine (NH_2CH_2COOH) a deux atomes de carbone. Ayant marqué au carbone 14 le carbone du CH_2, il retrouvait ce carbone dans la porphyrine extraite du sang du canard. Quand il marquait le carbone du COOH, il ne retrouvait pas le traceur radioactif dans la porhyrine. En résumé, le groupement CH_2 entre dans la synthèse de la porphyrine, contrairement au COOH.

Shemin, alors qu'il travaillait avec Rittenberg, trouva que l'incorporation des atomes de la glycine dans la porphyrine pouvait avoir lieu aussi bien dans les globules rouges et dans le tube à essai que dans l'animal vivant. Cette découverte simplifia la question, procura des résultats plus clairs, et évita de sacrifier ou de martyriser des animaux.

Shemin marqua alors l'azote de la glycine avec de l'azote 15, et le carbone de son CH_2 avec du carbone 14, puis mélangea la glycine à du sang de canard. Il dégrada ensuite avec soin la porphyrine formée, et trouva que les quatre atomes d'azote de la molécule de porphyrine provenaient de la glycine ; il en est de même de l'atome de carbone adjacent, dans chaque petit cycle pyrrolique (voir la formule au chapitre 11), et des quatre atomes servant de ponts entre ces cycles. On compte encore douze atomes de

carbone dans le cycle porphyrique lui-même, et quatorze dans les chaînes latérales variées. On a montré qu'ils proviennent des ions acétates, certains du carbone du CH_3, et d'autres du COO_-.

De la distribution des atomes marqués, on a pu déduire la façon dont la glycine et l'acétate participent à la porphyrine. Un cycle pyrrolique unique est d'abord formé ; ensuite deux de ces cycles s'associent, et finalement ces derniers s'unissent pour former la structure à quatre cycles de la porphyrine.

En 1952, un composé, le *porphobilinogène*, fut isolé sous forme pure, selon une ligne de recherche indépendante suivie par le chimiste anglais R.G. Westall. On trouve ce composé dans l'urine de personnes ayant des troubles du métabolisme de la porphyrine ; on soupçonna donc que ce composé avait quelque chose à voir avec les porphyrines. Il s'avéra que sa structure était identique à celle du cycle pyrrolique unique que Shemin et ses collaborateurs supposaient être une étape précoce de la synthèse de la porphyrine. Le porphobilinogène en est une étape clef.

On montra ensuite que l'*acide delta-aminolévulinique*, une substance ayant la structure d'un cycle de porphobilinogène coupé en deux, peut fournir tous les atomes nécessaires à la synthèse du cycle porphyrique par les globules rouges. L'explication la plus plausible est la suivante : les cellules commencent par former l'acide delta-aminolévulinique à partir de la glycine et de l'acétate (en éliminant le groupement COOH de la glycine sous forme de gaz carbonique au cours du processus) ; puis deux molécules d'acide delta-aminolévulinique se combinent pour former le porphobilinogène (un cycle pyrrolique unique) ; ce dernier s'associe d'abord en un cycle à deux pyrroles, et finalement en un cycle à quatre pyrroles.

La photosynthèse

C'est sans doute l'un des plus grands succès remportés par les chercheurs travaillant au moyen de traceurs ; il s'agit de comprendre la succession complexe d'étapes qui permet l'existence des plantes vertes, source de la vie sur notre planète.

Le règne animal ne pourrait pas exister si les animaux ne se nourrissaient qu'entre eux. Un lion qui mange un zèbre ou un homme qui mange un steak consomment une substance très précieuse qui a été obtenue au prix de grands efforts et par une contribution importante du monde végétal. Le deuxième principe de la thermodynamique nous dit qu'on perd quelque chose à chaque étape du cycle. Aucun animal ne conserve tous les glucides, graisses et protéines contenus dans la nourriture qu'il mange ; il ne peut pas non plus utiliser toute l'énergie disponible dans la nourriture. Inévitablement, une grande partie, en fait la plus grande partie de cette énergie est perdue sous forme de chaleur inutilisable. A chaque étape de l'action de manger, de l'énergie est perdue. Donc, si tous les animaux étaient rigoureusement des carnivores, le règne animal dans son entier mourrait en quelques générations. En réalité, il n'aurait jamais pu exister.

La chance veut que la plupart des animaux soient herbivores. Ils se nourrissent de l'herbe des prés, des feuilles des arbres, de graines, de noix et de fruits, ou d'algues et de plantes vertes microscopiques, dont les

couches supérieures de l'Océan sont pleines. Seule une minorité d'animaux peuvent se permettre le luxe d'être carnivores.

Quant aux plantes, elles ne seraient pas logées à meilleure enseigne, si on ne leur fournissait pas une source d'énergie extérieure. Elles synthétisent des glucides, des graisses et des protéines à partir de molécules simples, telles que le gaz carbonique et l'eau. Cette synthèse exige un apport d'énergie, et les plantes l'obtiennent de la source la plus abondante qui soit : le soleil. Les plantes vertes convertissent l'énergie solaire en énergie chimique sous forme de composés complexes, et cette énergie chimique maintient toutes les formes de vie (à l'exception de certaines bactéries). Ce processus a été clairement mis en évidence, en 1845, par le physicien allemand Julius Robert von Mayer, qui fut un des pionniers en matière d'économies d'énergie ; il était donc particulièrement au fait du problème de l'équilibre de l'énergie. Le mécanisme par lequel les plantes vertes utilisent le soleil s'appelle la *photosynthèse,* du grec signifiant « assemblé par la lumière ».

LE PROCESSUS

Les premières recherches sur la croissance des plantes ont été faites au début du XVIIe siècle par le chimiste flamand Jan Baptista Van Helmont. Il fit pousser un saule dans un baquet contenant une quantité déterminée de terre, et découvrit à la surprise de tous, au bout d'un certain temps, que quoique l'arbre ait bien poussé, le poids de la terre était inchangé. Il était admis alors que les plantes tirent leur substance du sol. En fait, elles utilisent quelques minéraux et ions du sol, mais en quantités qui ne sont pas aisément mesurables. D'où peuvent-elles donc bien tirer leur substance ? Van Helmont pensa que les plantes devaient fabriquer leur substance à partir de l'eau, qu'il leur avait fournie généreusement. Il n'avait que partiellement raison.

Au siècle suivant, le physiologiste anglais Stephen Hales montra que les plantes tirent leur substance essentiellement d'un matériau plus éthéré que l'eau, à savoir l'air. Un demi-siècle plus tard, le médecin hollandais Jan Ingen-Housz identifia le gaz carbonique comme étant l'ingrédient nourricier de l'air. Il démontra aussi que les plantes n'absorbent pas de gaz carbonique dans l'obscurité ; elles ont besoin de lumière (le *photo* de photosynthèse). Au même moment Priestley, celui qui a découvert l'oxygène, affirmait que les plantes libèrent de l'oxygène. Et, en 1804, le chimiste suisse Nicolas Théodore de Saussure prouva que l'eau fait partie intégrante du tissu des plantes, comme l'avait suggéré Van Helmont.

Dans les années 1850, l'ingénieur des mines français Jean-Baptiste Boussingault apporta une contribution importante en cultivant des plantes sur des sols ne contenant aucune matière organique. Il apporta ainsi la preuve que la seule source de carbone des plantes est le gaz carbonique de l'air. D'autre part, les plantes ne peuvent pas pousser sur des sols sans azote : elles tirent donc leur azote du sol et non de l'azote atmosphérique (ce que font certaines bactéries). Avec Boussingault, il devenait évident que le rôle du sol était limité à la fourniture de sels minéraux, tels que les nitrates et les phosphates, dans l'alimentation des plantes. Ce sont ces éléments qui sont apportés au sol par les engrais organiques tels que le fumier. Les chimistes préconisèrent l'addition d'engrais chimiques, qui

remplissent parfaitement leur tâche tout en éliminant les odeurs désagréables et en réduisant les dangers d'infections et de maladies.

Le schéma de la photosynthèse était ainsi établi. Au soleil, la plante absorbe le gaz carbonique, le lie à l'eau pour former ses tissus, et libère les « restes » d'oxygène au cours du processus. Les plantes fournissent donc de la nourriture et elles renouvellent aussi les réserves d'oxygène de la Terre. Sans ce renouvellement, le taux d'oxygène tomberait à un faible niveau en quelques siècles, et la concentration de gaz carbonique dans l'atmosphère serait si élevée, que la vie animale en serait asphyxiée.

L'échelle à laquelle les plantes vertes fabriquent de la matière organique et libèrent de l'oxygène est énorme. Le biochimiste américain d'origine russe Eugene I. Rabinowitch, un des chercheurs de premier plan de la photosynthèse, estime que chaque année les plantes vertes de la planète combinent un total de 150 milliards de tonnes de carbone (du gaz carbonique) à 25 milliards de tonnes d'hydrogène (de l'eau) et libèrent 400 milliards de tonnes d'oxygène. Dans cette performance gigantesque, les plantes des forêts et des champs n'entrent que pour 10 % ; les 90 % restants sont fournis par les plantes unicellulaires et les algues des océans.

LA CHLOROPHYLLE

Nous ne connaissons encore que les grandes lignes du processus. Qu'en est-il des détails ? En 1817, les Français Pierre Joseph Pelletier et Joseph Bienaimé Caventou, qui découvrirent plus tard la quinine, la strychnine, la caféine ainsi que d'autres produits végétaux, isolèrent le produit végétal le plus important de tous, celui qui donne sa couleur verte aux plantes : la *chlorophylle,* du mot grec signifiant « feuille verte ». Puis en 1865, le botaniste allemand Julius von Sachs montra que la chlorophylle n'est pas distribuée uniformément dans la cellule végétale (bien que les feuilles apparaissent uniformément vertes), mais qu'elle est localisée dans des corpuscules subcellulaires, les *chloroplastes.*

La photosynthèse se passe dans les chloroplastes et la chlorophylle est essentielle au processus. Cependant celle-ci n'est pas suffisante ; elle ne peut pas seule, même préparée soigneusement, catalyser la réaction photosynthétique dans un tube à essai.

Les chloroplastes sont généralement considérés comme de grandes mitochondries. Certaines plantes unicellulaires ne contiennent qu'un gros chloroplaste par cellule. La plupart des cellules végétales contiennent cependant jusqu'à quarante petits chloroplastes, chacun étant deux ou trois fois plus gros qu'une mitochondrie typique.

La structure du chloroplaste semble être encore plus complexe que celle de la mitochondrie. L'intérieur d'un chloroplaste est constitué de nombreuses membranes fines allant d'une paroi à l'autre. Ce sont les *lamelles.* Dans la plupart des espèces de chloroplastes, ces lamelles s'épaississent et s'assombrissent par endroits pour donner le *granum,* où l'on trouve les molécules de chlorophylle.

Les grana de la lamelle, vus au microscope, semblent constitués d'unités minuscules, à peine visibles, qui ressemblent aux carreaux bien alignés du sol d'une salle de bains. Ces unités de photosynthèse peuvent contenir de 250 à 300 molécules de chlorophylle.

Il est plus difficile d'isoler des chloroplastes intacts que des mitochondries. Ce n'est qu'en 1954 que le biochimiste américain d'origine polonaise,

Daniel Israel Arnon, en travaillant avec des cellules de feuilles d'épinards broyées, a obtenu des chloroplastes parfaitement intacts et a pu mener à bien une réaction complète de la photosynthèse.

En plus de la chlorophylle, le chloroplaste contient un jeu complet d'enzymes et de substances associées, le tout organisé de façon complexe. Il contient même des cytochromes, qui permettent à l'énergie solaire d'être captée par la chlorophylle et convertie en ATP grâce à l'oxydation phosphorylante.

La chlorophylle est la substance la plus caractéristique des chloroplastes ; mais quelle est sa structure ? Pendant des décades, les chimistes se sont attaqués à cette substance importante, avec tous les moyens à leur portée, mais elle ne s'est rendue que lentement. Finalement, en 1906, l'Allemand Richard Willstätter (qui découvrit plus tard la chromatographie et insista sur le fait que les enzymes ne sont pas des protéines) identifia le composant central de la molécule de chlorophylle, le magnésium. Willstätter reçut le prix Nobel de chimie en 1915 pour ses travaux sur les pigments des plantes. Willstätter et Hans Fischer travaillèrent à la structure de la molécule, une tâche qui occupa toute une génération. En 1930, on savait déjà que la chlorophylle contient un noyau porphyrique de structure semblable à celui de l'hème, une molécule que Fischer avait décrite. Là où l'hème a un atome de fer, au centre du noyau prophyrique, la chlorophylle a un atome de magnésium.

Tout doute à ce sujet fut éliminé par R.B. Woodward. Ce maître de la synthèse – qui a produit la quinine en 1945, la strychnine en 1947, et le cholestérol en 1951 – couronna alors ses efforts précédents en synthétisant, en 1960, une molécule se conformant à la formule établie par Willstätter et Fischer. Celle-ci avait en plus toutes les propriétés de la chlorophylle isolée des feuilles vertes. Woodward reçut le prix Nobel de chimie en 1965.

Quelle réaction la chlorophylle catalyse-t-elle exactement ? La seule chose que l'on savait en 1930, c'était que le gaz carbonique et l'eau entrent dans la cellule tandis que l'oxygène en sort. Les recherches étaient rendues plus difficiles par le fait que la chlorophylle purifiée ne peut provoquer la photosynthèse. Seules les cellules de plantes intactes, ou au mieux des chloroplastes isolés en sont capables ; le système à étudier est donc très complexe.

A première vue, les biochimistes supposèrent que les cellules des plantes synthétisent le glucose ($C_6H_{12}0_6$) à partir du gaz carbonique et de l'eau, puisqu'elles s'en servent pour assembler leurs différents éléments, en ajoutant de l'azote, du soufre, du phosphore, et les autres composants minéraux tirés du sol.

Sur le papier, il semble que le glucose soit polymérisé à partir du formaldéhyde (CH_2O), qui serait formé au cours d'une série d'étapes où le carbone du gaz carbonique s'associe d'abord à l'eau (libérant les atomes d'oxygène du CO_2), six molécules de formaldéhyde pouvant former une molécule de glucose.

On peut synthétiser le glucose dans un tube à essai, à partir du formaldéhyde, mais c'est très fastidieux. Les plantes possèdent vraisemblablement des enzymes qui accélèrent les réactions. Le formaldéhyde est un composé très toxique, mais les chimistes pensaient qu'il se transforme si vite en glucose que les plantes n'en contiennent jamais que des traces. Cette théorie du formaldéhyde, qui fut proposée en 1870 par Baeyer (celui

qui a synthétisé l'indigo), a tenu bon pendant deux générations parce qu'il n'y avait simplement rien de mieux pour la remplacer.

En 1938, Ruben et Kamen s'attaquèrent au problème sous un autre angle : ils explorèrent la chimie des feuilles vertes à l'aide de traceurs. En utilisant l'oxygène 18, l'isotope rare et stable de l'oxygène, ils firent une découverte claire et nette : quand une plante est arrosée avec de l'eau marquée à l'oxygène 18, elle libère de l'oxygène marqué. Au contraire, quand on fournit à la plante du gaz carbonique marqué, l'oxygène formé par la plante n'est pas marqué. Autrement dit, l'expérience montrait que l'oxygène libéré par les plantes provient de la molécule d'eau et non pas du gaz carbonique, comme cela avait été supposé par erreur dans la théorie du formaldéhyde.

Ruben et ses associés tentèrent de suivre le devenir des atomes de carbone dans la plante en marquant le gaz carbonique par le carbone 11 radioactif (le seul disponible à cette époque). Ils échouèrent, parce que la demi-vie du carbone 11 n'est que de vingt minutes et demie, et qu'ils ne disposaient pas de méthode de séparation des différents constituants de la cellule végétale qui soit assez rapide ni complète.

Mais au début des années quarante, les outils nécessaires devinrent disponibles. En effet, Ruben et Kamen découvrirent le carbone 14, l'isotope ayant une longue vie, qui leur permit de suivre le carbone tout au long d'une série de réactions. La mise au point de la chromatographie sur papier leur procura un moyen de séparer facilement et avec une bonne résolution des mélanges complexes. En fait, les radio-isotopes permirent une amélioration sensible de la chromatographie sur papier : la présence du traceur crée des taches radioactives sur le papier, qui forment des taches sombres sur un film photographique mis à son contact, de sorte que le chromatogramme se prend lui-même en photo ; cette technique est appelée *autoradiographie.*

Après la Deuxième Guerre mondiale, un autre groupe, dirigé par le biochimiste américain Melvin Calvin, prit le relais. Ces chercheurs exposèrent des plantes unicellulaires microscopiques (*chlorella*) à du gaz carbonique contenant du carbone 14, pendant des durées de temps courtes, pour permettre à la photosynthèse de n'accomplir que les étapes les plus précoces du processus. Ils broyèrent ensuite les cellules des plantes, en séparèrent le contenu par chromatographie, et firent un autoradiogramme.

Ils découvrirent que, même lorsque les cellules n'avaient été exposées au gaz carbonique marqué que pendant une minute et demie, ils retrouvaient dans la cellule le carbone radioactif associé à au moins quinze substances différentes. En diminuant le temps d'exposition, ils réduisirent le nombre de substances dans lesquelles le carbone radioactif était incorporé, et finalement conclurent que le premier composé dans lequel le gaz carbonique était incorporé était le *glycérol phosphate.* Ils ne détectèrent jamais la moindre trace de formaldéhyde. Cette vénérable théorie s'évanouit.

Le glycérol phosphate est un composé à trois carbones. Il doit être formé par une voie indirecte, puisqu'aucun composé précurseur à un ou deux carbones n'a pu être détecté. Deux autres composés incorporant le carbone marqué très rapidement furent mis en évidence. Ce sont deux sucres : le *ribulose diphosphate* (un composé à cinq carbones) et le *sédoheptulose phosphate* (un composé à sept carbones). Les chercheurs identifièrent les enzymes qui catalysent les réactions impliquant ces sucres, étudièrent ces

réactions, et déterminèrent le parcours suivi par le gaz carbonique, qui peut être résumé de la façon suivante.

Le gaz carbonique est d'abord fixé sur le ribulose diphosphate à cinq carbones, constituant un composé à six carbones. Celui-ci est rapidement coupé en deux, créant ainsi le glycérol phosphate à trois carbones. Une série de réactions, impliquant le sédoheptulose phosphate et d'autres composés, permettent la condensation de deux molécules de glycérol phosphate pour former le glucose phosphate à six carbones. Le ribulose phosphate est régénéré en même temps, et prêt à accepter une nouvelle molécule de gaz carbonique. On peut imaginer ainsi six cycles tournants, qui à chaque tour de cycle fournissent chacun un atome de carbone (provenant du gaz carbonique), et à partir de ces six carbones, une molécule de glucose est construite. Un autre tour de six cycles produit une autre molécule de glucose phosphate, etc.

Du point de vue énergétique, c'est l'inverse du cycle tricarboxylique. Alors que le cycle tricarboxylique convertit les fragments obtenus dans la dégradation des glucides en gaz carbonique, le cycle du ribulose diphosphate fabrique des glucides à partir du gaz carbonique. Le cycle tricarboxylique fournit de l'énergie à l'organisme ; inversement, le cycle du ribulose diphosphate consomme de l'énergie.

C'est ici que s'intègrent les résultats de Ruben et Kamen. L'énergie solaire est utilisée, grâce à l'action catalytique de la chlorophylle, pour scinder une molécule d'eau en hydrogène et oxygène ; c'est la *photolyse* (du grec signifiant « décomposition par la lumière »). C'est la façon dont l'énergie rayonnante du soleil est convertie en énergie chimique, car les molécules d'hydrogène et d'oxygène contiennent plus d'énergie que la molécule d'eau dont elles proviennent.

Sans la photolyse, il faut beaucoup d'énergie pour casser une molécule d'eau, par exemple en chauffant l'eau à quelque 2 000 °C ou en y faisant passer un fort courant électrique. La chlorophylle peut cependant scinder facilement une molécule d'eau à la température ordinaire. Tout ce dont elle a besoin, c'est de l'énergie relativement faible de la lumière visible. La plante utilise l'énergie de la lumière qu'elle absorbe avec une efficacité d'au moins 30 % ; certains chercheurs pensent que son efficacité peut atteindre 100 % dans des conditions idéales. Si nous, êtres humains, pouvions domestiquer l'énergie avec autant d'efficacité que les plantes, nous serions bien moins préoccupés de nos réserves de nourriture et d'énergie.

La molécule d'eau ayant été scindée, la moitié des atomes d'hydrogène entrent dans le cycle du ribulose diphosphate, et la moitié des atomes d'oxygène sont libérés dans l'air. Le reste des oxygènes et des hydrogènes se réassocient pour former de l'eau. Ce faisant, ils dégagent l'excès d'énergie reçu quand le soleil a divisé les molécules d'eau en leurs éléments. Cette énergie est transférée aux composés phosphorylés à haute énergie, tels que l'ATP. L'énergie conservée dans ces composés est alors utilisée pour alimenter le cycle du ribulose diphosphate. Calvin reçut le prix Nobel de chimie en 1961 pour avoir décrypté les réactions impliquées dans la photosynthèse.

Il existe, sans aucun doute, des formes de vie qui acquièrent de l'énergie sans chlorophylle. Aux environs de 1880, on découvrit les *bactéries chimiosynthétiques ;* ce sont des bactéries qui, dans l'obscurité, piègent le gaz carbonique, mais ne libèrent pas d'oxygène. Certaines oxydent des

composés soufrés pour acquérir de l'énergie, d'autres oxydent des composés ferreux, d'autres encore s'adonnent à d'autres fantaisies chimiques.

Il existe aussi des bactéries dont les composés, similaires à la chlorophylle (les *bactériochlorophylles*), leur permettent de convertir le gaz carbonique en composés organiques aux dépens de l'énergie de la lumière, utilisant dans certains cas l'infrarouge, dans lequel la chlorophylle ordinaire n'est pas fonctionnelle. Cependant, seule la chlorophylle peut scinder la molécule d'eau et conserver les grandes quantités d'énergie ainsi produites ; les bactériochlorophylles doivent se débrouiller avec des systèmes énergétiques moins efficaces.

Tous les moyens permettant l'acquisition d'énergie fondamentale, autres que ceux utilisant la lumière solaire à l'aide de la chlorophylle, sont essentiellement des culs de sac, et n'existent que dans des conditions complexes, rares et particulières. Seules les bactéries sont capables de s'en servir avec succès. Pour à peu près tout ce qui vit, la chlorophylle et la photosynthèse, directement ou indirectement, sont le fondement de la vie.

Chapitre 13

La cellule

Les chromosomes

Il peut paraître étrange que, récemment encore, nous ne connaissions pas grand-chose de notre propre corps. En fait, nous n'avons commencé à comprendre la circulation du sang qu'il y a trois siècles, et ce n'est que durant ce dernier demi-siècle que nous avons découvert la fonction de nombreux organes.

Les hommes préhistoriques avaient pris conscience de l'existence des organes principaux, tels que le cerveau, le foie, le cœur, les poumons, l'estomac, les intestins et les reins, en découpant le gibier pour leur nourriture, et en embaumant leurs morts pour l'au-delà. Cette prise de conscience s'accentua lorsqu'on se mit à utiliser les organes internes des animaux sacrifiés selon des rituels (en particulier le foie) pour prédire l'avenir ou évaluer la faveur ou la défaveur divines. C'est ainsi que des papyrus égyptiens datant d'avant 2000 av. J.-C. traitent de techniques chirurgicales qui laissent supposer une certaine connaissance de la structure de l'organisme.

Les anciens Grecs allèrent jusqu'à disséquer des animaux, et occasionnellement des cadavres humains, dans le but d'apprendre l'*anatomie* (du grec signifiant « découper »). Des travaux précis furent accomplis. Alcméon, environ en 500 av. J.-C., fut le premier à décrire le nerf optique et la trompe d'Eustache. Deux siècles plus tard, à Alexandrie, en Égypte (qui était alors le centre du monde scientifique), une école d'anatomie grecque débuta brillamment avec Hérophile et son élève Érasistrate. Ils firent des

recherches sur certaines parties du cerveau, établissant la distinction entre le cerveau et le cervelet, et étudièrent aussi les nerfs et les vaisseaux.

L'anatomie des Anciens atteignit son apogée lorsque Galien, un médecin grec qui pratiquait à Rome pendant la deuxième moitié du IIe siècle, établit des théories au sujet des fonctions corporelles, qui eurent valeur de dogme pendant les quinze siècles qui suivirent. Cependant, ses idées sur le corps humain comportaient de curieuses erreurs, aisément compréhensibles, puisque les Anciens tiraient leurs informations de la dissection des animaux. L'idée de disséquer le corps humain mettait les gens mal à l'aise.

Les chrétiens qualifièrent les Grecs de païens et les accusèrent d'avoir pratiqué des vivisections sur les êtres humains. Mais ceci appartient à la littérature polémique ; non seulement cette affirmation est douteuse, mais encore les Grecs n'avaient pas disséqué suffisamment de cadavres pour en apprendre assez sur l'anatomie humaine. Dans tous les cas, la désapprobation de la dissection par l'Église jugula la recherche anatomique pendant le Moyen Age. Vers la fin de cette période, l'anatomie commença à ressusciter en Italie. En 1316, un anatomiste italien, Mondino de Luzzi, écrivit le premier livre entièrement consacré à l'anatomie ; il est par conséquent considéré comme le « restaurateur de l'anatomie ».

L'intérêt porté à l'art naturaliste pendant la Renaissance stimula la recherche anatomique. Au XVe siècle, Léonard de Vinci pratiqua quelques dissections qui lui permirent de révéler des faits anatomiques nouveaux, qu'il dessina avec toute sa puissance artistique. Il démontra la double courbure de la colonne vertébrale ainsi que le sinus creusé dans les os de la face et du front. Il utilisa ses recherches pour formuler des théories plus avancées que celles de Galien. Mais Vinci, quoiqu'un génie en science et en art, n'a guère eu d'influence sur la pensée scientifique de son temps. Il ne publia aucun de ses travaux scientifiques, qu'il conserva cachés dans des carnets de notes codés. Ce n'est que quelques générations plus tard que ses découvertes scientifiques ont été rendues publiques, lorsque ses carnets de notes furent finalement publiés.

Le médecin français Jean Fernel fut le premier homme moderne à considérer l'anatomie comme partie intégrante des fonctions du médecin. Il publia un livre à ce sujet en 1542. Mais son travail fut presque complètement éclipsé par une contribution beaucoup plus importante qui parut l'année suivante. C'était le célèbre *De corporis humani fabrica libri septem* (de la structure du corps humain) de André Vésale, un Belge qui réalisa presque tous ses travaux en Italie. Vésale partait du principe que pour étudier convenablement l'homme, il fallait le disséquer, et c'est ce qu'il fit. Il corrigea ainsi bon nombre d'erreurs de Galien. Les dessins d'anatomie humaine illustrant son livre (qui auraient été faits par Jan Stevenzoon Van Calcar, un élève du Titien) sont si beaux et si exacts qu'ils sont encore reproduits de nos jours et resteront toujours des classiques. On peut considérer que Vésale est le père de l'anatomie moderne. Sa *Fabrica* a été aussi révolutionnaire que le *De revolutionibus orbium cœlestium* de Copernic, publié la même année.

Tout comme la révolution commencée par Copernic fut amenée à maturité par Galilée, celle amorcée par Vésale arriva à point nommé avec les découvertes cruciales de William Harvey. Harvey était un expérimentateur et médecin anglais, de la même génération que Galilée et William Gilbert (l'expérimentateur du magnétisme). Le centre d'intérêt de Harvey était le fluide vital de l'organisme, le sang. Quelle était donc sa fonction dans le corps ?

On savait qu'il existe deux sortes de vaisseaux : les veines et les artères. Praxagoras de Cos, un médecin grec du IIIe siècle av. J.-C., avait donné le nom d'*artères*, du grec signifiant « transporter de l'air », à ces vaisseaux trouvés vides dans les cadavres. Galien montra plus tard que ces vaisseaux transportent du sang dans le corps vivant. On savait aussi que les battements du cœur propulsent le sang d'une certaine façon car, lorsque les artères sont coupées, le sang en jaillit avec des pulsions qui sont synchronisées avec les battements cardiaques.

Galien avait suggéré que le sang allait et venait dans les vaisseaux sanguins, se déplaçant à travers le corps, d'abord dans une direction, puis dans l'autre. Cette théorie l'obligeait à expliquer pourquoi le mouvement de va-et-vient du sang n'était pas bloqué par la paroi entre les deux moitiés du cœur ; Galien répondit simplement que la paroi était criblée de petits trous invisibles qui laissaient passer le sang.

Harvey examina le cœur de plus près. Il trouva que chaque moitié était divisée en deux cavités séparées par une valvule à sens unique, permettant au sang de s'écouler de la cavité supérieure (*l'oreillette*) vers la cavité inférieure (*le ventricule*), mais pas inversement. En d'autres termes, le sang entrant dans l'une des oreillettes s'écoule dans le ventricule correspondant, et de là dans les vaisseaux sanguins qui en partent, mais il ne peut y avoir de circulation dans le sens opposé.

Harvey fit alors des expériences simples mais claires pour déterminer le sens de la circulation dans les vaisseaux sanguins. Il ligatura des veines et des artères d'animaux vivants, et regarda de quel côté de ce barrage le sang s'accumulait. Il trouva que s'il bloquait une artère, le vaisseau se gonflait toujours entre le cœur et la ligature. Quand il ligaturait une veine, le gonflement était toujours de l'autre côté de l'obstacle. Donc le sang circulant dans les artères s'éloigne du cœur, alors que celui des veines s'y rend. Le fait que les plus grosses veines contiennent des valvules qui empêchent le sang de s'éloigner du cœur constitue une preuve supplémentaire de cette circulation à sens unique. Ce mécanisme a été découvert par le professeur de Harvey, l'anatomiste italien Hieronymus Fabrizzi (plus connu sous son nom latinisé, Fabricius). Mais comme Fabricius était sous l'emprise de la tradition de Galien, il refusa d'en tirer les conclusions et laissa la gloire à son étudiant.

Harvey mesura alors le flux sanguin (c'était la première fois que l'on appliquait les mathématiques à un problème biologique). Ses mesures montrèrent que le cœur éjecte le sang à une vitesse telle qu'il débite en vingt minutes la quantité totale du sang contenue dans l'organisme. Il ne semblait pas raisonnable de penser que l'organisme fabriquait du sang neuf, ou consommait le vieux, à un tel rythme. La conclusion logique était donc que le sang devait se recycler dans le corps. Puisqu'il circule à partir du cœur dans les artères et vers le cœur dans les veines, le sang est pompé par le cœur dans les artères, puis il passe dans les veines, et retourne au cœur, où il est pompé à nouveau dans les artères, etc. En d'autres termes, il circule sans arrêt dans la même direction à l'intérieur du système circulatoire.

Auparavant, d'autres anatomistes, dont Léonard de Vinci, avaient suggéré cette idée, mais Harvey fut le premier à la formuler et à l'étudier en détail. Il a exposé son raisonnement et ses expériences dans un petit livre, mal imprimé, intitulé *De motû cordis* (du mouvement du cœur), publié en 1628, qui est considéré depuis comme un des grands classiques scientifiques.

Il restait la question importante à laquelle les travaux de Harvey n'apportaient pas de réponse : comment le sang passe-t-il des artères aux veines ? Harvey supposa qu'il devait exister des vaisseaux pour cela, trop petits pour être vus. Cette hypothèse rappelait la théorie de Galien à propos des trous dans la paroi du cœur ; mais tandis que les trous imaginés par Galien n'ont jamais pu être mis en évidence et n'ont jamais existé, les vaisseaux de liaison de Harvey furent confirmés dès l'invention du microscope. En 1661, quatre ans après la mort de Harvey, un médecin italien, Marcello Malpighi, examina les tissus des poumons d'une grenouille à l'aide d'un microscope primitif, et observa en effet que des vaisseaux sanguins minuscules reliaient les artères aux veines. Malpighi les nomma *capillaires,* du latin signifiant « comme un cheveu ». (Pour le schéma du système circulatoire, voir figure 13.1.)

L'utilisation du microscope permit aussi de voir d'autres structures minuscules. Le naturaliste hollandais Jan Swammerdam découvrit les globules rouges, tandis que l'anatomiste hollandais Régnier de Graaf découvrit les minuscules follicules ovariens dans les ovaires animaux. De petites créatures, telles que les insectes, pouvaient enfin être étudiées avec précision.

L'étude de si petits détails encouragea la comparaison des structures d'une espèce à l'autre. Le botaniste anglais Nehemiah Grew fut le premier à pratiquer l'*anatomie comparée* et à se faire une notoriété dans le domaine. En 1675, il publia ses études sur la comparaison de la structure du tronc de différents arbres, et en 1681 ses études sur la comparaison des estomacs de différents animaux.

LA THÉORIE CELLULAIRE

L'arrivée du microscope amena les biologistes à un niveau d'organisation plus fondamental des organismes vivants, celui où toutes les structures ordinaires peuvent être réduites à un dénominateur commun. En 1665, le savant anglais Robert Hooke, utilisant un microscope composé qu'il avait conçu, découvrit que le liège, l'écorce du chêne, était constitué de compartiments minuscules à la manière d'une éponge très fine ; il les appela *cellules,* parce qu'ils ressemblaient à de petites pièces, comme les cellules d'un monastère. D'autres utilisateurs du microscope trouvèrent ensuite des cellules similaires dans les tissus vivants, mais remplies de fluide.

Durant le siècle et demi qui suivit, les biologistes commencèrent progressivement à entrevoir que toute matière vivante est constituée de cellules, et que chaque cellule est une unité de vie indépendante. Certaines formes de vie – les micro-organismes – sont constituées d'une seule cellule ; les organismes plus grands sont composés de nombreuses cellules qui coopèrent. La proposition la plus précoce dans ce sens émana du physiologiste français René Joachim Henri Dutrochet. Son mémoire, publié en 1824, passa inaperçu ; et la théorie cellulaire ne prit de l'importance qu'après que Mathias Jacob Schleiden et Theodor Schwann, tous les deux allemands, l'eurent formulée séparément en 1838 et 1839.

Le fluide colloïdal qui emplit certaines cellules fut appelé *protoplasme* (« forme primitive ») par le physiologiste tchèque Jan Evangelista Purkinje, en 1839, et le botaniste allemand Hugo von Mohl étendit le terme au contenu de toutes les cellules. L'anatomiste allemand Max Johann

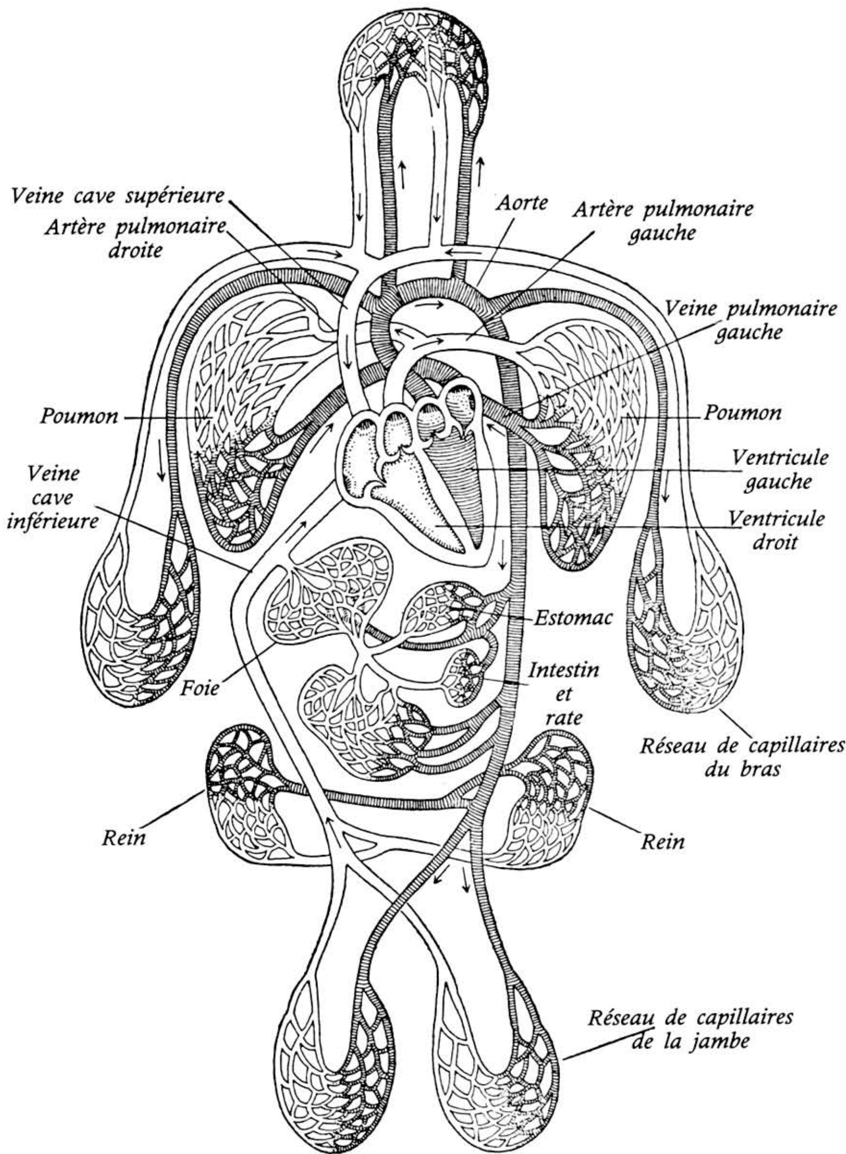


Figure 13.1. Le système circulatoire.

Sigismund Schultze mit en valeur l'importance du protoplasme en tant que « base physique de la vie », et démontra la similitude du protoplasme chez toutes les cellules, aussi bien végétales qu'animales, aussi bien chez les créatures simples que chez les êtres complexes.

La théorie cellulaire est à la biologie ce que la théorie de l'atome est à la physique et à la chimie. Son importance dans la dynamique de la

vie fut établie quand, aux environs de 1860, le pathologiste allemand Rudolf Virchow affirma, en une formule latine succincte, que toutes les cellules proviennent de cellules. Il montra que les cellules des tissus malades sont produites par la division de cellules originellement saines.

A cette époque, il était devenu évident que tous les organismes vivants, même les plus gros, sont à l'origine une seule et unique cellule. Un des premiers microscopistes, Johann Ham, assistant de Leeuwenhoek, avait découvert des corpuscules minuscules dans le fluide séminal, qui furent appelés plus tard *spermatozoïdes*, du grec signifiant « semence animale ». Beaucoup plus tard, en 1827, le physiologiste allemand Karl Ernst von Baer identifia l'*ovule*, ou cellule de l'œuf, des mammifères (figure 13.2).

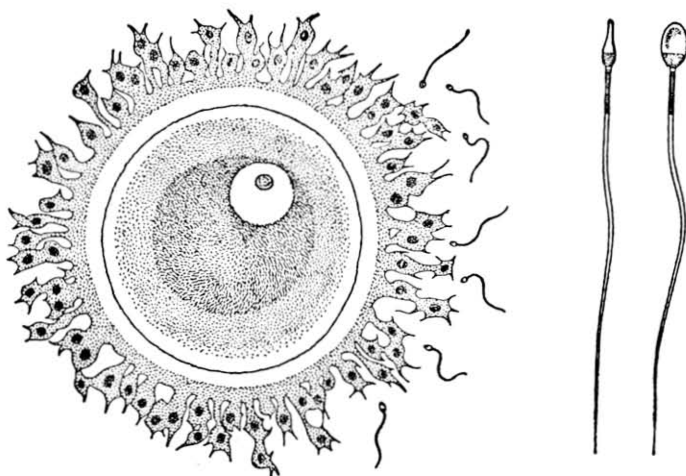


Figure 13.2. L'ovule et le spermatozoïde humains.

Les biologistes comprirent que l'union d'un œuf et d'un spermatozoïde forme un ovule fertilisé d'où l'animal se développe par une succession de divisions et de subdivisions.

Les plus grands organismes n'ont donc pas des cellules plus grandes que les petits organismes : ils en ont simplement davantage. Les cellules demeurent petites, presque toujours microscopiques. La cellule typique de la plante ou de l'animal a un diamètre compris entre cinq et quarante micromètres (un *micromètre*, ou *micron* est un millionième de mètre), et l'œil humain peut à peine distinguer un objet ayant cent micromètres de diamètre.

Bien que les cellules soient si petites, ce ne sont en aucun cas des gouttelettes de protoplasme informes. La cellule a une infrastructure complexe qui a été décryptée petit à petit, à partir du XIXe siècle. C'est vers cette infrastructure que les biologistes se sont tournés pour répondre aux nombreuses questions concernant la vie.

Par exemple, étant donné que la croissance des organismes est due à la multiplication des cellules qui les constituent, comment les cellules se divisent-elles ? La réponse réside dans de petits corpuscules de densité

optique relativement forte, situés à l'intérieur de la cellule, et occupant environ un dixième de son volume. C'est Robert Brown (à l'origine de la découverte du mouvement brownien) qui les observa le premier et qui les appela *noyaux* (pour les distinguer du noyau de l'atome, je m'y référerai en tant que *noyaux cellulaires*).

Si on coupe un organisme unicellulaire en deux parties, celle contenant le noyau cellulaire dans son entier croît et se divise tandis que l'autre ne le peut pas. On découvrit plus tard que les globules rouges du sang des mammifères qui ne contiennent pas de noyau ont une vie courte et ne peuvent ni croître ni se diviser. C'est pour cette raison qu'ils ne sont pas considérés comme de véritables cellules et qu'ils sont appelés *globules*.

Malheureusement, l'étude plus approfondie du noyau cellulaire et du mécanisme de la division fut contrariée pendant longtemps parce que la cellule est plus ou moins transparente, et que son infrastructure n'est pas visible. La situation s'améliora quand on découvrit que certains colorants ne teignent que certaines parties de la cellule. Un colorant, l'*hématoxaline* (obtenue du bois), teint le noyau en violet et le fait ressortir, par contraste, du reste de la cellule. Après que Perkin et les autres chimistes eurent produit des colorants synthétiques, les biologistes se retrouvèrent avec un grand choix de colorants.

En 1879, le biologiste allemand Walther Flemming trouva que certains colorants rouges teignaient un élément du noyau cellulaire distribué sous forme de petits granules, qu'il appela *chromatine* (du mot grec pour « couleur »). En l'examinant, Flemming suivit les changements qui s'opéraient au cours de la division cellulaire. Évidemment, le colorant tuait la cellule, mais dans une coupe de tissu, on pouvait voir les cellules à différents stades de la division cellulaire. C'était comme des photographies, qu'il assembla dans l'ordre exact pour construire un « film » de la progression de la cellule lors d'une division cellulaire.

En 1882, Flemming publia un livre important où il décrivit le processus en détail. Au début de la division cellulaire, la chromatine se condense sous la forme de filaments. La fine membrane entourant le noyau cellulaire semble se dissoudre ; en même temps un petit objet, à l'extérieur du noyau, se divise en deux. Flemming l'appela *aster*, du mot grec « étoile », parce que ses filaments en forme de rayons lui donnent l'aspect d'une étoile. Après s'être divisées, les deux parties de l'aster se dirigent vers les deux pôles opposés de la cellule. Les filaments semblent s'emmêler avec ceux de la chromatine, qui entre-temps s'est alignée au centre de la cellule, et l'aster attire la moitié de la chromatine vers un pôle de la cellule et l'autre moitié vers l'autre pôle. Il en résulte que la cellule se pince en son milieu et se scinde en deux cellules. Un noyau cellulaire se reforme dans chacune d'elles, et la chromatine se rassemble à nouveau en granules (voir la figure 13.3).

Flemming appela le processus de la division cellulaire *mitose*, du mot grec « filament », la chromatine y jouant un rôle important. En 1888, l'anatomiste allemand Wilhelm von Waldeyer donna le nom de *chromosomes* aux filaments de chromatine (du grec « corps coloré »), et le nom leur est resté. Il faut cependant mentionner que malgré leur nom, les chromosomes sont incolores à l'état naturel, dans lequel il est alors difficile de les distinguer du reste de la cellule. Pourtant, le botaniste amateur allemand Wilhelm Friedrich Benedict Hofmeister les avait aperçus dans les cellules de fleurs dès 1848.

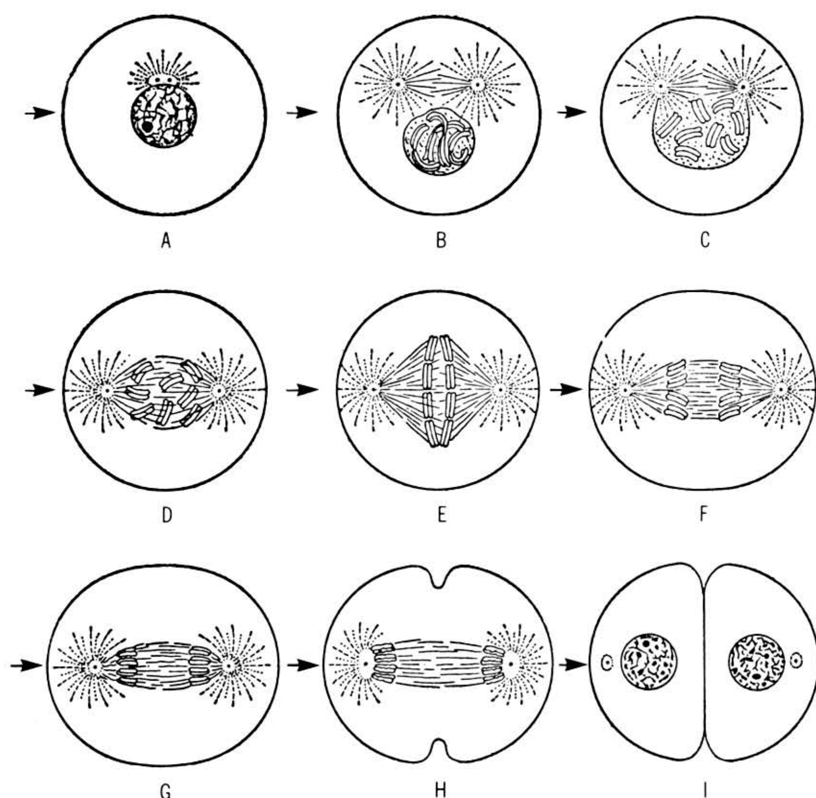


Figure 13.3. La division cellulaire par mitose.

La poursuite de l'observation de cellules colorées a montré que chaque espèce végétale ou animale a un nombre caractéristique et fixe de chromosomes. Pendant la mitose, avant que la cellule ne se divise en deux, le nombre de chromosomes double, de sorte que chaque cellule fille, après la division, a le même nombre de chromosomes que sa mère.

L'embryologiste belge Édouard Van Beneden découvrit en 1885 que le nombre de chromosomes *ne double pas* lors de la formation de l'ovule ou du spermatozoïde. Chaque ovule ou spermatozoïde n'a donc que la moitié du nombre de chromosomes que possèdent les autres cellules de l'organisme. La division cellulaire conduisant à la formation des spermatozoïdes et des ovules est appelée *méiose,* du grec « décroissance ». Quand l'œuf et le spermatozoïde fusionnent, la cellule résultante (l'œuf fécondé) a un jeu complet de chromosomes, la moitié provenant de la mère par l'ovule, et l'autre moitié du père par le spermatozoïde. Ce jeu complet est transmis ensuite à toutes les cellules constituant l'organisme qui se développe à partir de l'œuf, par simple mitose.

Bien que l'utilisation de colorants permette de voir les chromosomes, elle ne permet pas de les identifier individuellement. La plupart du temps, ils ressemblent à un enchevêtrement de gros spaghettis. On a donc pensé pendant très longtemps que les cellules humaines contenaient vingt-quatre paires de chromosomes ; et ce n'est qu'en 1956 qu'un comptage plus fin montra que le nombre correct est de vingt-trois.

Heureusement, ce problème est résolu. On a mis au point une technique qui consiste à faire gonfler des cellules par traitement, dans des conditions données, à l'aide d'une solution de faible concentration saline, qui permet de disperser les chromosomes. On peut alors les photographier, découper sur les photographies les chromosomes séparés, et les ranger par taille décroissante. On obtient un *caryotype*, c'est-à-dire une image du contenu chromosomique de la cellule dont les chromosomes sont numérotés en séquence.

Le caryotype est un outil très précis de diagnostic médical, car la séparation des chromosomes n'est pas toujours parfaite. Au cours de la division cellulaire, un chromosome peut être endommagé ou même cassé. Quelquefois la séparation est inégale, de sorte que l'une des cellules filles se retrouve avec un chromosome supplémentaire, tandis qu'il en manque un à l'autre. De telles anomalies, évidemment, modifient le fonctionnement de la cellule, et peuvent même l'entraver totalement ; ce qui participe au maintien de la fidélité apparente de la mitose.

Ces imperfections sont particulièrement dramatiques quand elles ont lieu au cours de la méiose, car l'ovule ou le spermatozoïde produit possède alors un jeu de chromosomes anormal. Si l'organisme réussit à se développer malgré tout sur cette base (ce qui en général n'est pas possible), chaque cellule du corps sera imparfaite ; il en résultera une maladie congénitale grave.

La maladie la plus fréquente se traduit par une arriération mentale. C'est le *syndrome de Down* (parce qu'il a été décrit en premier par le médecin anglais John Langdon Haydon Down, en 1866), dont la fréquence est de un pour mille naissances. Il est plus connu sous le nom de *mongolisme*, car l'un de ses symptômes consiste en une inclinaison des paupières qui rappelle les yeux bridés des peuples d'Extrême-Orient. Mais cette appellation ne se justifie pas vraiment, car le syndrome a la même fréquence chez les Asiatiques que dans les autres ethnies.

La cause du syndrome de Down n'a été découverte qu'en 1959. Cette année-là trois généticiens français, Jérôme Jean Lejeune, Marthe Gautier et Raymond Turpin, comptèrent les chromosomes des cellules de trois personnes présentant le syndrome, et en trouvèrent quarante-sept, au lieu de quarante-six, dans chaque cas. Il s'est avéré que l'erreur résidait dans la présence de *trois* chromosomes 21. La maladie symétrique a été détectée en 1967. Une petite fille de trois ans, retardée mentale, n'avait qu'un chromosome 21. C'est le premier être humain connu vivant avec un chromosome en moins.

D'autres cas similaires, impliquant d'autres chromosomes, semblent moins fréquents, mais il y en a. Chez les malades présentant un certain type de leucémie, on note un échange de fragments entre deux chromosomes, qui crée la *translocation de Philadelphie*, détectée en premier chez un patient de cette ville. On trouve des chromosomes cassés avec une fréquence supérieure à la moyenne dans certaines maladies peu communes.

LA REPRODUCTION ASEXUÉE

La formation d'un nouvel individu à partir d'un œuf fécondé dont la moitié des chromosomes provient de chaque parent est la *reproduction sexuée*. C'est la norme pour l'être humain et en général pour les organismes qui ont notre niveau de complexité.

La *reproduction asexuée* existe cependant, le nouvel individu ne possédant que l'assortiment de chromosomes d'un seul parent. Un exemple de reproduction asexuée est le cas de l'être unicellulaire qui se divise en deux, formant deux cellules indépendantes ayant chacune le même jeu de chromosomes que la cellule mère.

La reproduction asexuée est très fréquente aussi dans le monde végétal. La tige d'une plante peut être plantée dans le sol où elle prend racine et croît, produisant un organisme complet semblable à celui d'où provient la tige. Ou bien une brindille peut être greffée sur la branche d'un autre arbre (même d'une autre variété) où elle peut croître et fleurir. On appelle cette brindille un *clone*, du mot grec « pousse » ; le terme *clone* est utilisé maintenant communément pour désigner tous les organismes d'origine asexuée ayant un seul parent.

La reproduction asexuée peut aussi avoir lieu chez des êtres pluricellulaires. Les animaux les plus primitifs – c'est-à-dire ceux dont les cellules sont les moins diversifiées et spécialisées – sont ceux où la reproduction asexuée a le plus de chances d'avoir lieu.

Une éponge, une hydre d'eau douce, un ver plat, une étoile de mer, peuvent être mis en morceaux, et si ces morceaux sont conservés dans leur environnement habituel, chacun reformera un organisme complet. Les nouveaux organismes sont considérés comme des clones.

Même des organismes aussi complexes que les insectes peuvent, dans certains cas, donner naissance à des descendants provenant d'un seul parent. Les pucerons (aphidiens), par exemple, le font tout naturellement. Dans leur cas, un œuf non fertilisé, ne contenant qu'un demi-jeu de chromosomes, peut se débrouiller sans spermatozoïde. Le demi-jeu de chromosomes de l'œuf se dédouble, produisant un assortiment complet entièrement d'origine maternelle ; l'œuf peut alors se diviser et devenir un organisme indépendant, un clone en quelque sorte.

Cependant, chez les animaux complexes, il ne se produit généralement pas de clonage naturel, et la reproduction y est exclusivement sexuée. L'intervention de l'homme peut toutefois provoquer le clonage des vertébrés.

Un œuf fécondé est capable de produire un organisme complet ; cet œuf se divise et se redivise, chaque cellule nouvelle contenant un jeu complet de chromosomes identiques au jeu de départ. Pourquoi donc chacune des nouvelles cellules ne posséderait-elle pas la capacité de produire un nouvel individu si on l'isole et la conserve dans des conditions qui permettent à l'œuf fertilisé de se développer ?

Alors que l'œuf fécondé se divise et se redivise, les nouvelles cellules se *différencient*, devenant des cellules hépatiques, cutanées, nerveuses, musculaires, rénales, etc. Chacune de ces cellules a des fonctions très différentes de celles des autres ; et il est vraisemblable que les chromosomes subissent des modifications qui rendent possible cette différenciation cellulaire. Ce sont ces changements subtils qui font que la cellule différenciée est incapable de former un nouvel individu à partir de rien.

Mais les chromosomes sont-ils modifiés d'une façon permanente et irréversible ? Que se passerait-il si on remettait ces chromosomes dans l'environnement de l'œuf ? Supposons, par exemple, que l'on isole l'ovule non fertilisé d'un animal d'une espèce donnée, et que l'on élimine le noyau avec soin. Supposons aussi que l'on prépare de la même façon le noyau d'une cellule de la peau de ce même individu, et qu'on l'introduise dans l'œuf. Sous l'influence de l'ovule qui est conçu pour assurer la croissance d'un individu complet, les chromosomes du noyau de la cellule de la peau, qui ne remplissent qu'une partie de leurs fonctions, ne devraient-ils pas retrouver leur pluripotentialité originelle ? Est-ce que l'œuf, *fécondé* ainsi, pourra se développer pour produire un nouvel individu, ayant le même jeu de chromosomes que l'individu à qui appartenaient les cellules de la peau donneuses du noyau fécondateur ? Le nouvel individu ainsi obtenu n'est-il pas un clone du donneur des cellules de peau ?

Naturellement, l'élimination du noyau et son remplacement sont des opérations extrêmement délicates, mais elles ont été menées à bonne fin, en 1952, par les biologistes américains Robert William Briggs et Thomas J. King. Leurs travaux sont à la base de la technique de *transplantation nucléaire*.

En 1967, le biologiste britannique John B. Gordon a transplanté le noyau d'une cellule de l'intestin d'un crapaud d'Afrique du Sud dans un œuf non fertilisé de la même espèce, et de cet œuf un individu parfaitement normal est né – une première dans le domaine du clonage.

Répéter cette expérience avec des œufs de reptiles ou d'oiseaux serait très difficile, car ils sont entourés d'une coquille, et on ne peut les conserver vivants et fonctionnels après que la coquille a été ouverte pour faire pénétrer le noyau.

Qu'en est-il des œufs de mammifères ? Ils sont nus, mais sont conservés à l'intérieur du corps de la mère ; ils sont particulièrement petits et fragiles, et il faudrait utiliser des techniques de microchirurgie, qui ont encore besoin d'être améliorées.

On a pu, cependant, faire des transplantations nucléaires chez la souris ; et, en principe, le clonage devrait être possible chez tous les mammifères, y compris chez l'homme.

Les gènes

LA THÉORIE MENDÉLIENNE

Dans les années 1860, un moine autrichien, Gregor Johann Mendel, trop occupé par les affaires de son monastère pour prêter attention à l'enthousiasme des biologistes au sujet de la division cellulaire, poursuivait tranquillement des expériences dans son jardin, destinées à comprendre la signification des chromosomes. L'abbé Mendel, botaniste amateur, s'intéressa en particulier aux résultats de croisements entre des plants de pois présentant des caractères variés. Son grand coup de génie intuitif a été de n'étudier qu'un seul caractère bien défini à la fois.

Il a croisé des plants ayant des graines de couleurs différentes, d'autres ayant des graines lisses avec des plants à graine plissée, des plants à tige

longue avec des plants à tige courte, et en a suivi les résultats sur plusieurs générations. Mendel consignait soigneusement les statistiques obtenues et ses conclusions peuvent être résumées ainsi :

1 Chaque caractère est gouverné par des *facteurs* qui (dans les cas étudiés par Mendel) ne peuvent exister que sous l'une des deux formes possibles. Par exemple, une version du facteur pour la couleur de la graine fera que la graine sera verte ; tandis que l'autre forme la rendra jaune. Pour plus de facilité, utilisons le terme *gène,* qui a remplacé celui de facteur de nos jours. Le mot gène a été proposé en 1909 par le biologiste danois Wilhelm Ludwig Johannsen ; il vient du mot grec signifiant « naissance » ; et les différentes formes d'un gène contrôlant un caractère défini sont appelées *allèles*. Les gènes de la couleur de la graine possèdent donc deux allèles, un pour les graines vertes, et l'autre pour les graines jaunes.

2 Les différents caractères d'une plante sont contrôlés par une paire de gènes dont l'un provient du mâle et l'autre de la femelle. La plante transmet un des gènes de sa paire à une cellule germinale (un terme général pour désigner le spermatozoïde et l'ovule), de sorte que lorsque deux cellules germinales s'unissent par la pollinisation, la progéniture possède de nouveau deux gènes pour ce caractère. Les deux gènes peuvent être identiques ou allèles.

3 Quand les deux parents transmettent à leurs descendants les deux allèles d'un gène donné, un allèle peut dominer l'effet de l'autre. Par exemple, si on croise une plante à graines jaunes avec une plante à graines vertes, tous les descendants de la première génération auront des graines jaunes. L'allèle jaune du gène gouvernant la couleur de la graine est *dominant,* l'allèle vert est *récessif.*

4 L'allèle récessif n'est cependant pas détruit. L'allèle de la couleur verte, dans le cas mentionné, est toujours présent, bien qu'il ne produise aucun effet détectable. Si on croise deux plantes contenant les deux allèles (c'est-à-dire chacune ayant un allèle jaune et un allèle vert), une partie de leur descendance aura deux gènes verts dans l'ovule fertilisé ; ces descendants-là produiront des graines vertes, et la descendance de tels parents produira aussi des graines vertes. Mendel a fait remarquer qu'il y a quatre façons différentes d'associer les deux allèles de parents hybrides (ayant les deux allèles : jaune et vert). L'allèle jaune d'un parent peut s'associer à l'allèle jaune de l'autre ou à son allèle vert ; et l'allèle vert du premier parent peut s'associer soit à l'allèle jaune, soit à l'allèle vert du second. De ces quatre combinaisons, seule la dernière donnera une plante à graines vertes. Si les quatre combinaisons sont également probables, un quart des plantes de la nouvelle génération devrait produire des graines vertes. Ce que Mendel observa.

5 Mendel trouva aussi que des caractères de nature différente, la couleur de la graine et celle de la fleur, par exemple, peuvent être héritées indépendamment les unes des autres, c'est-à-dire que l'on peut retrouver des fleurs rouges aussi bien avec des graines vertes que des graines jaunes. C'est aussi vrai des fleurs blanches.

Mendel réalisa ses expériences au début des années 1860, les rédigea avec soin et envoya une copie de son mémoire à Karl Wilhelm von Nägeli, un botaniste suisse de grande réputation. La réaction de celui-ci fut négative ; il avait apparemment une préférence pour les théories globales. Il ne trouva aucun mérite au fait de compter simplement les plants de pois pour trouver la vérité. De plus, Mendel n'était qu'un obscur amateur.

Il semble que les critiques de von Nägeli aient affecté Mendel, car il retourna à ses devoirs monastiques, grossit au point de ne plus pouvoir se baisser dans le jardin, et abandonna ses recherches. Il publia quand même son article dans un journal provincial autrichien, en 1866 ; personne n'y prêta attention pendant une génération.

D'autres savants arrivèrent lentement aux mêmes conclusions que Mendel, dont ils ne connaissaient rien. Une des voies par laquelle ils aboutirent à la génétique fut l'étude des *mutations*, à savoir des monstres qui ont toujours été considérés comme des mauvais présages. Le mot *monstre* provient d'un mot latin signifiant « alarme ». En 1791, Seth Wright, un fermier du Massachusetts, réagit avec pragmatisme à une anomalie qui survint dans son troupeau de moutons. Un agneau naquit avec des pattes anormalement courtes, et l'astucieux yankee pensa qu'un mouton à pattes courtes ne pouvait pas sauter le mur de la ferme. Il sélectionna donc délibérément une lignée de moutons à pattes courtes, mettant à profit l'incident.

Cette démonstration pratique encouragea d'autres gens à rechercher des mutations utiles. A la fin du XIXe siècle, l'horticulteur américain Luther Burbank poursuivit une carrière prospère en sélectionnant des centaines de nouvelles variétés de plantes qui étaient des améliorations des anciennes, obtenues par mutations, croisements ou greffes.

Les botanistes, pendant ce temps-là, essayaient de trouver une explication aux mutations. Et par la coïncidence la plus étonnante de l'histoire de la science, trois hommes parvinrent séparément, et la même année, aux mêmes conclusions précisément que Mendel, une génération plus tôt. C'étaient Hugo de Vries en Hollande, Karl Erich Correns en Allemagne et Erich von Tschermak en Autriche. Aucun d'eux ne connaissait les travaux des autres, ni ceux de Mendel ; tous les trois étaient prêts à publier en 1900 ; chacun, lors d'une dernière vérification des publications antérieures dans la spécialité, découvrit avec étonnement l'article de Mendel. Ils publièrent tous les trois en 1900, chacun citant cet article, donnant tout le mérite de la découverte à Mendel, et ne présentant son propre travail que comme une confirmation des travaux de celui-ci.

LE PATRIMOINE GÉNÉTIQUE

Certains biologistes firent immédiatement le lien entre les gènes de Mendel et les chromosomes vus au microscope. Le premier à établir ce parallèle fut Walter Stanborough Sutton, un cytologiste, en 1904. Il fit remarquer que les chromosomes, comme les gènes, sont appariés ; un membre de la paire provenant de la mère, l'autre étant hérité du père. Le seul ennui dans cette analogie, c'est que les chromosomes présents dans la cellule sont beaucoup moins nombreux que les caractères héréditaires. L'homme, par exemple, ne possède que vingt-trois paires de chromosomes, alors qu'il a sûrement des milliers de caractères héréditaires. La conclusion des biologistes fut que les chromosomes ne sont pas des gènes. Chaque chromosome est une collection de gènes.

Les biologistes découvrirent très rapidement un excellent outil pour étudier des gènes spécifiques. Ce n'était pas un nouvel instrument de physique, mais un nouvel animal de laboratoire. En 1906, le zoologiste

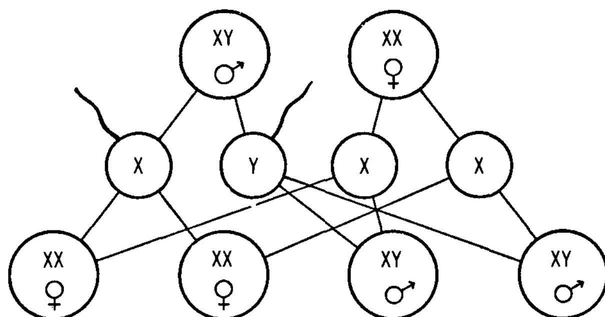
Thomas Hunt Morgan, de l'université Columbia à New York, qui était de prime abord sceptique au sujet de la théorie de Mendel, conçut l'idée d'utiliser la mouche du vinaigre (*Drosophila melanogaster*) pour la recherche en *génétique* (ce terme fut inventé par le biologiste anglais William Bateson en 1902).

La mouche du vinaigre présentait des avantages considérables par rapport aux pois (ou aux animaux de laboratoire habituels) pour étudier la transmission des gènes : elle se reproduit rapidement et de façon prolifique, elle peut aisément être élevée par centaines avec très peu de nourriture, elle a des tas de caractères héréditaires qui sont facilement observables et elle n'a que quatre paires de chromosomes, ce qui est très peu.

Grâce à la mouche du vinaigre, Morgan et ses collaborateurs découvrirent un fait important concernant la transmission du sexe. Ils trouvèrent que la femelle de la mouche du vinaigre a quatre paires de chromosomes parfaitement appariées, de telle sorte que tous les ovules, recevant chacun une paire de chaque chromosome, sont identiques en ce qui concerne leur composition chromosomique. Chez le mâle, il n'en est pas de même ; l'une des quatre paires de chromosomes est constituée d'un chromosome normal, *le chromosome X*, et d'un chétif, le *chromosome Y*. Quand les spermatozoïdes sont formés, la moitié reçoivent donc un chromosome X, et l'autre moitié un chromosome Y. Quand un spermatozoïde portant le chromosome X féconde un ovule, l'œuf fertilisé, ayant quatre paires de chromosomes assortis, deviendra naturellement une femelle. En revanche, un spermatozoïde porteur du chromosome Y produira un mâle. Étant donné que l'alternative est également probable, le nombre de mâles et de femelles dans une population typique d'êtres vivants est virtuellement le même (figure 13.4). Chez certains êtres, notamment certains oiseaux, c'est la femelle qui possède le chromosome Y.

Cette différence chromosomique explique pourquoi certaines mutations ou maladies ne touchent que les mâles. S'il arrive qu'un gène défectueux se trouve sur un chromosome X, l'autre gène sur le deuxième chromosome X peut être normal et sauver ainsi la situation. Mais chez le mâle, un défaut sur le chromosome X, apparié au chromosome Y, ne peut pas toujours être compensé, car il manque quelques gènes à ce dernier, et l'anomalie devient apparente.

Figure 13.4. L'association des chromosomes X et Y.



L'exemple le plus connu de *maladie liée au sexe* est l'*hémophilie*, dans laquelle le sang se coagule avec difficulté ou pas du tout. Les individus hémophiles courent le risque de saigner à mort pour des égratignures ou de souffrir le martyre en cas d'hémorragie interne. La femme possédant deux chromosomes X a de bonnes chances de n'avoir qu'un seul gène défectueux, l'autre étant normal ; elle est *porteuse* du gène de l'hémophilie et pourtant elle n'est pas malade. La moitié de ses ovules possédera le chromosome normal et l'autre moitié le chromosome X hémophile. Si l'ovule ayant le chromosome X anormal est fécondé par un spermatozoïde d'un mâle normal comportant un chromosome X, la fille qui en résultera ne sera pas hémophile, mais sera aussi porteuse ; si l'ovule est fécondé par un spermatozoïde ayant un chromosome Y, provenant d'un mâle normal, le gène hémophile de l'ovule ne sera pas compensé, et le fils qui naîtra sera hémophile. Selon la loi du hasard, les fils d'une porteuse de l'hémophilie ont une chance sur deux d'être hémophiles et la moitié des filles seront à leur tour porteuses.

La porteuse de l'hémophilie la plus connue de l'histoire est la reine Victoria d'Angleterre. Un seul de ses quatre fils (Léopold, l'aîné) fut hémophile. Édouard VII – dont descendent les souverains actuels – y échappa ; il n'y a donc plus d'hémophiles dans la famille royale. Cependant, deux des filles de Victoria étaient aussi porteuses. L'une d'entre elles eut une fille (porteuse elle aussi) qui épousa le tzar Nicolas II de Russie. C'est pourquoi son unique fils fut hémophile ; l'histoire de la Russie et du monde s'en trouva modifiée, car c'est par son influence sur ce fils hémophile que le moine Grégoire Raspoutine acquit un certain pouvoir en Russie, provoquant le mécontentement qui, en fin de compte, conduisit à la révolution. L'autre fille de Victoria eut aussi une fille porteuse qui épousa un membre de la famille royale d'Espagne, et y introduisit des hémophiles. A cause de sa présence chez les Bourbons d'Espagne et chez les Romanoff de Russie, l'hémophilie est quelquefois appelée *maladie royale* ; elle n'a cependant rien à voir avec la royauté, en dehors de l'infortune de Victoria.

Une maladie qui a une liaison moins importante avec le sexe est le daltonisme, bien plus fréquent chez l'homme que chez la femme. L'absence d'un chromosome X contribue à diminuer la longévité masculine et explique en partie pourquoi dans les pays où les morts par accouchements sont rares, les femmes vivent de trois à sept ans de plus en moyenne que les hommes. La vingt-troisième paire complète de chromosomes fait, en quelque sorte, de la femme l'organisme biologique le plus solide. On a suggéré récemment que la mortalité plus grande de l'homme était due au fait qu'il fume, mais depuis que les femmes fument aussi et que les hommes fument moins, leur taux de mortalité tend à s'égaliser.

Les chromosomes X et Y (ou les *chromosomes sexuels*) sont placés arbitrairement à la fin du caryotype, bien que le chromosome X soit parmi les plus longs. Les anomalies chromosomiques sont apparemment plus fréquentes sur les chromosomes sexuels que sur les autres. La raison n'est peut-être pas que les chromosomes sexuels sont plus souvent impliqués dans des mitoses anormales, mais que leurs anomalies sont moins fatales, de sorte que plus d'enfants arrivent à terme malgré tout.

L'anomalie touchant les chromosomes sexuels qui a le plus attiré l'attention est celle où le mâle se retrouve avec un chromosome Y supplémentaire dans ses cellules ; il est donc XYY. Il s'avère que les mâles

XYY sont de caractère difficile. Ils sont grands, forts et vifs, mais ils manifestent aussi une tendance prononcée à la violence. On soupçonne Richard Speck, qui tua huit infirmières à Chicago en 1966, d'être XYY. Un meurtrier fut acquitté en Australie en octobre 1968, parce qu'il était XYY, et donc jugé non responsable de ses actes. Près de 4 % des mâles internés dans une prison écossaise se sont révélés être XYY, alors que l'on estime que seulement un mâle sur 3 000 est porteur de l'association XYY.

Il serait souhaitable, semble-t-il, d'effectuer un contrôle chromosomique sur chaque individu et en tout cas sur tous les nouveau-nés. Comme c'est le cas de nombreuses techniques, simples en théorie mais longues et laborieuses en pratique, on étudie actuellement la possibilité de le faire faire par l'ordinateur.

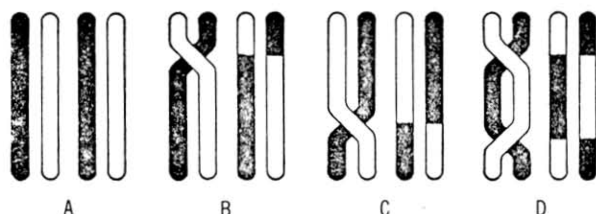
LA RECOMBINAISON

Les recherches sur la mouche du vinaigre ont montré que les caractères ne sont pas forcément hérités indépendamment, comme Mendel l'avait pensé. Les sept caractères des pois qu'il avait étudiés étaient contrôlés par des gènes situés sur des chromosomes différents. Morgan a découvert que si deux gènes gouvernant deux caractères différents sont situés sur le même chromosome, ces caractères sont généralement transmis en même temps (comme les passagers avant et arrière d'une voiture voyagent ensemble).

La liaison génétique est modifiable. De même que le passager peut changer de place, un morceau de chromosome peut occasionnellement s'échanger avec un autre morceau de chromosome. Une telle *recombinaison* peut se produire pendant la division cellulaire (figure 13.5). Il en résulte que les caractères qui étaient liés sont séparés et réorganisés dans une nouvelle liaison. Par exemple, il existe une variété de mouche du vinaigre aux yeux vermillon et aux ailes frisées. Quand on la croise avec une mouche à yeux blancs et ailes miniatures, la descendance a généralement soit des yeux vermillon et des ailes frisées, soit des yeux blancs et des ailes miniatures. Mais le croisement peut quelquefois, à cause d'une recombinaison, donner naissance à des mouches à yeux vermillon et petites ailes ou bien à yeux blancs et ailes frisées. Cette nouvelle variété persistera au cours des générations successives, sauf si une nouvelle recombinaison intervient.

Imaginons un chromosome portant le gène des yeux vermillon à une extrémité et le gène des ailes frisées à l'autre. Supposons qu'en plus, au milieu du chromosome, il y ait deux gènes adjacents gouvernant deux

Figure 13.5. La recombinaison chromosomique.



autres caractères. Il est évident qu'une rupture du chromosome a plus de chances de se produire entre les deux gènes éloignés qu'entre les deux gènes adjacents. En notant la fréquence avec laquelle deux caractères liés sont séparés par recombinaison, Morgan et ses collaborateurs, notamment Alfred Henry Sturtevant, purent localiser les gènes les uns par rapport aux autres et établir ainsi une *carte chromosomique* donnant la localisation des gènes de la mouche du vinaigre. L'emplacement d'un gène ainsi déterminé est le *locus* du gène.

Mais comme cela arrive souvent quand on étudie des systèmes biologiques, leur comportement ne suit pas toujours de manière rigide les règles comme les scientifiques aiment à le penser. A partir des années quarante, la biologiste américaine Barbara McClintock a étudié avec soin les gènes du maïs et les a suivis de génération en génération ; elle est arrivée à la conclusion que, au cours de la division cellulaire, certains gènes se déplacent fréquemment et facilement le long du chromosome. Cette constatation sortait tellement du schéma établi par Morgan et les biologistes qui le suivirent qu'on ignora ses travaux, mais elle avait raison. Tandis que d'autres chercheurs établissaient la preuve de la mobilité de certains gènes, Barbara McClintock (maintenant octogénaire) reçut le prix Nobel de physiologie et de médecine en 1983.

En se servant des cartes chromosomiques et de l'étude des chromosomes géants (que l'on trouve dans la glande salivaire de la mouche du vinaigre, et qui sont beaucoup plus grands que les chromosomes normaux), on a déterminé qu'il y a un minimum de 10 000 gènes dans une paire de chromosomes de la mouche du vinaigre. Un gène individuel doit donc avoir un poids moléculaire de 60 000 000. En conséquence, les chromosomes humains, qui sont plus grands, peuvent contenir de 20 000 à 90 000 gènes par paire, soit au moins 2 000 000 de gènes en tout.

Morgan reçut le prix Nobel de physiologie et de médecine en 1933 pour ses travaux sur la mouche du vinaigre.

La connaissance accrue des gènes donne l'espoir que la dotation génétique de chaque homme pourra un jour être analysée et modifiée, soit en empêchant des anomalies graves de se développer, soit en les corrigeant si elles apparaissent malgré tout. Un tel *génie génétique* exige des cartes des chromosomes humains ; une tâche bien plus considérable que dans le cas de la mouche du vinaigre. La tâche a été simplifiée de façon étonnante par Howard Green, de l'université de New York. Il a construit des cellules hybrides contenant à la fois des chromosomes humains et des chromosomes de souris. Étant donné qu'il restait relativement peu de chromosomes humains dans les cellules après plusieurs divisions, les effets dus à leur activité ont pu être plus facilement mis en évidence.

Un nouveau pas dans la connaissance des gènes et de leur manipulation a été fait en 1969, quand le biochimiste américain Jonathan Beckwith et ses collaborateurs ont isolé un gène pour la première fois dans l'histoire. Il a été isolé d'une bactérie de l'intestin, et il contrôle une étape du métabolisme du sucre.

LE FARDEAU GÉNÉTIQUE

De temps en temps, avec une fréquence qui peut être calculée, un gène change brusquement. La *mutation* est révélée par une caractéristique

physique nouvelle et inattendue, telle que les pattes courtes de l'agneau de Wright, le fermier. Les mutations sont relativement rares dans la nature. En 1926, Herman Joseph Muller, un généticien, ancien membre de l'équipe de Morgan, découvrit un moyen artificiel d'augmenter la fréquence des mutations chez la drosophile, en utilisant les rayons X qui endommagent les gènes. On put ainsi étudier plus facilement la transmission des modifications provoquées. Cette étude valut à Muller le prix Nobel de médecine et de physiologie en 1946.

Comme c'est souvent le cas, les recherches de Muller ont soulevé des questions quelque peu inquiétantes au sujet de l'avenir de l'espèce humaine. Bien que les mutations soient une importante force agissante de l'évolution, produisant occasionnellement une amélioration qui permet à l'espèce de mieux s'adapter à son environnement, les mutations bénéfiques sont exceptionnelles. La plupart des mutations, 99 % d'entre elles, sont défavorables, certaines sont même létales. En fin de compte, celles qui ne sont que légèrement nuisibles finissent par disparaître, parce que ceux qu'elles atteignent se portent moins bien et laissent une descendance plus faible que les individus sains. Mais en attendant, une mutation cause des maladies et des souffrances pendant de nombreuses générations. De plus, de nouvelles mutations continuent à s'accumuler et chaque espèce est porteuse d'un fardeau constant de tares. En fait, plus de 1 600 maladies humaines sont imputables à des tares génétiques.

Le généticien américain d'origine russe Theodosius Dobzhansky montra, dans les années trente et quarante, qu'une population normale possède une grande variété de gènes, y compris des gènes très nocifs. Cette diversité permet à l'évolution de fonctionner comme elle le fait, mais le nombre de gènes délétères (le *fardeau génétique*) est à l'origine d'une anxiété justifiée.

Deux évolutions modernes semblent alourdir régulièrement ce fardeau. D'abord, les progrès de la médecine et une meilleure prise en charge sociale tendent à compenser les handicaps des gens ayant des mutations défavorables, au moins tant que la capacité de se reproduire n'est pas concernée. Les individus ayant une vision défectueuse disposent de lunettes ; l'insuline permet aux diabétiques (une maladie héréditaire) de survivre, etc. Ils transmettent donc leurs gènes défectueux aux générations suivantes. L'alternative – laisser mourir les jeunes handicapés, les stériliser, ou les emprisonner – est bien sûr impensable, sauf quand le handicap est suffisamment grand pour que ces individus ne soient plus tout à fait des humains comme dans le cas de l'idiotie ou de la paranoïa homicide. Sans doute l'espèce humaine peut-elle encore supporter sa charge de gènes mutés négativement, malgré ses impulsions humanitaires.

Mais nous sommes moins excusables en ce qui concerne le deuxième danger moderne, à savoir l'accroissement des tares dues à une exposition inutile aux radiations. La recherche génétique montre de manière évidente que, pour une population entière, une légère augmentation de l'exposition globale aux radiations correspond à une augmentation du taux de mutations. Or, depuis 1895, nous sommes exposés à deux types de radiations en plus grande intensité que les générations précédentes, qui les ignoraient totalement. Le rayonnement solaire, la radioactivité naturelle du sol et les rayons cosmiques ont toujours été présents. Mais maintenant, nous utilisons sans compter les rayons X en médecine et chirurgie dentaire ; nous concentrons les matériaux radioactifs ; nous fabriquons des isotopes radioactifs artificiels ayant un rayonnement d'une

puissance terrifiante ; nous faisons même exploser des bombes atomiques. Tout cela augmente le niveau de base de rayonnement.

Personne ne suggère bien sûr que l'on abandonne la recherche en physique nucléaire ou que les rayons X ne soient plus utilisés par les médecins ou les dentistes. Il est cependant fortement recommandé de tenir compte de ce danger et de diminuer l'exposition aux rayonnements ; les rayons X devraient être utilisés avec discernement et avec soin, et les organes sexuels devraient être systématiquement protégés lors de leur utilisation. Une autre précaution à prendre serait que chaque individu note le nombre d'expositions aux rayons X auxquels il est exposé afin d'éviter de dépasser une limite raisonnable.

LES GROUPES SANGUINS

Pour les généticiens, les principes basés sur des expériences avec les plantes ou les insectes ne s'appliquaient pas obligatoirement à l'homme. Après tout, celui-ci n'est ni un pois, ni une mouche du vinaigre. Cependant, l'étude de certains caractères humains montre que la génétique humaine obéit aux mêmes règles. Le meilleur exemple en est la transmission des groupes sanguins.

La transfusion sanguine ne date pas d'aujourd'hui, et les premiers médecins essayèrent parfois de transfuser du sang animal à des malades affaiblis à la suite d'une perte de sang. Mais les transfusions, même de sang humain, se terminaient mal, au point que des lois furent édictées pour les interdire. Dans les années 1890, le pathologiste autrichien Karl Landsteiner découvrit enfin que le sang humain présente différents groupes, dont certains sont incompatibles entre eux. Il remarqua que parfois, quand le sang d'une personne est mélangé à un échantillon de *sérum* (le liquide sanguin restant après élimination des globules rouges et de facteurs de la coagulation) provenant d'une autre personne, les globules rouges du sang de la première personne s'agglutinent. Un tel mélange serait évidemment très dangereux s'il se produisait au cours d'une transfusion, pouvant même tuer le patient si les globules agglutinés bloquaient la circulation dans un gros vaisseau. Landsteiner observa aussi que certains sangs pouvaient être mélangés sans causer la moindre agglutination néfaste.

Dès 1902, Landsteiner fut en mesure d'annoncer l'existence de quatre groupes de sang humain, qu'il appela A, B, AB, et O. Chaque individu a un seul groupe sanguin. Bien entendu, un groupe donné peut être transfusé sans danger d'une personne à une autre du même groupe. De plus, le sang de groupe O peut être transfusé, sans inconvénient, à toute autre personne ayant n'importe quel autre groupe, et du sang de groupe A ou B peut être donné à une personne de groupe AB. Mais l'*agglutination* des globules rouges se produit quand on transfuse du sang AB à des receveurs A ou B, quand on mélange des sangs A et B, ou quand on transfuse un sang autre que du groupe O à un receveur O. De nos jours, à cause de réactions sériques possibles, il est recommandé de ne donner à un patient que le sang de son propre groupe.

En 1930, Landsteiner (qui, entre-temps, était devenu citoyen américain) reçut le prix Nobel de médecin et physiologie.

Les généticiens ont établi que ces groupes sanguins, ainsi que tous les autres découverts depuis, y compris le facteur rhésus (Rh), sont transmis

de manière strictement mendélienne. Il semble qu'il y ait trois gènes allèles responsables respectivement des groupes A, B et O. Si les deux parents sont de groupe O, tous leurs enfants seront de type O. Si un parent est de type O et l'autre de type A, les enfants peuvent être du groupe A, car celui-ci est dominant sur l'allèle O. De la même façon, l'allèle B est dominant sur O. Les allèles A et B ne présentent aucune dominance entre eux, et un individu possédant les deux allèles a un sang de groupe AB.

Les lois de Mendel gouvernent les groupes sanguins avec tant de précision qu'elles sont utilisées dans les tests de paternité. Si une mère de groupe O met au monde un enfant de groupe B, le père doit être de groupe B, car cet allèle doit venir de quelque part. Si le père est en fait A ou O, il est évident que la femme a été infidèle (ou qu'il y a eu confusion de bébé à l'hôpital). Si une femme de groupe O ayant un enfant de groupe B prétend qu'un homme de groupe A ou O en est le père, ou elle se trompe, ou elle ment. En revanche, si un groupe sanguin peut apporter une preuve négative, il ne peut pas apporter de preuve positive. Si le mari de la femme est vraiment un groupe B, le cas n'est pas résolu. N'importe quel homme de groupe B ou AB peut être le père.

L'EUGÉNISME

L'applicabilité des lois de Mendel aux êtres humains est aussi corroborée par l'existence de caractères liés au sexe. Comme je l'ai dit auparavant, le daltonisme et l'hémophilie se rencontrent presque exclusivement chez les mâles, et sont transmis exactement comme les caractères liés au sexe chez la mouche du vinaigre.

Naturellement, on peut penser que si on interdit aux sujets porteurs de telles infirmités d'avoir des enfants, celles-ci peuvent être éradiquées. En organisant correctement les accouplements, la race humaine pourrait même être améliorée, comme les races de bétail l'ont été. Ce n'est pas une idée nouvelle. Les Spartiates y croyaient et essayèrent de la mettre en pratique, il y a 2 500 ans. A l'époque moderne, elle a été reprise par un savant anglais, Francis Galton (un cousin de Charles Darwin) qui, en 1883, créa le terme *eugénisme* pour décrire le processus (le mot vient du grec et signifie « bonne naissance »).

Galton ignorait, à ce moment-là, les découvertes de Mendel. Il ne comprenait pas que certains caractères semblent absents et que, malgré tout, ils soient présents sous forme récessive. Il ne comprenait pas que des groupes de caractères soient transmis intacts, et qu'il soit dificile de se débarrasser d'un caractère indésirable sans se débarasser en même temps d'un caractère souhaitable. Il ne savait pas non plus qu'à chaque génération des mutations réintroduisent des tares dans la population.

Le désir « d'améliorer » la race humaine se perpétue néanmoins, et l'eugénisme a ses partisans, encore de nos jours, même parmi les chercheurs. Une telle attitude est presque toujours suspecte, car ceux qui cherchent à différencier d'un point de vue génétique les groupes d'êtres humains considéreront à coup sûr comme « supérieur » celui auquel ils appartiennent.

C'est ainsi que Cyril Lodowic Burt, un psychologue anglais, rendit compte d'études faites sur l'intelligence de différents groupes, et prétendit qu'il y avait de solides raisons de supposer que les hommes sont plus

intelligents que les femmes, que les chrétiens sont plus intelligents que les juifs, que les Anglais sont plus intelligents que les Irlandais, que les Anglais de haute extraction sont plus intelligents que ceux de milieu modeste, etc. Burt lui-même appartenait, dans chaque cas, au groupe « supérieur ». Ses résultats furent naturellement acceptés par de nombreuses personnes qui, comme lui, faisaient partie de la classe « supérieure », et qui étaient prêtes à croire que ceux qui s'en tiraient mal n'étaient pas les victimes de l'oppression ou d'un préjudice, mais de leur propres défauts.

Après la mort de Burt, en 1971, on se mit à douter de ses résultats. On se méfiait de la perfection de ses statistiques. Les soupçons grandirent ; et en 1978, le psychologue américain D.D. Dorfman montra, de façon concluante, que Burt avait simplement fabriqué ses résultats, tant il tenait à étayer la thèse en laquelle il croyait profondément, mais qui ne pouvait être démontrée honnêtement.

Malgré cela, Shockley, le co-inventeur du transistor, obtint une certaine notoriété en soutenant que les Noirs étaient génétiquement moins intelligents que les Blancs, de sorte qu'essayer d'améliorer leur sort en leur donnant des chances égales était voué à l'échec. Hans J. Eysenck, un psychologue germano-anglais, partage également ce point de vue.

En 1980, Shockley s'exposa à des plaisanteries de mauvais goût en révélant qu'il avait donné de ses spermatozoïdes, à soixante-dix ans, pour qu'ils soient conservés par congélation dans une banque de sperme, dans le but d'inséminer éventuellement une femme de grande intelligence qui le souhaiterait.

A mon avis la génétique humaine est un sujet trop compliqué pour se voir mené à bonne fin, complètement et correctement, dans un avenir proche. Parce que nous n'engendrons pas aussi souvent et de manière aussi prolifique que la mouche du vinaigre ; parce que nos croisements ne peuvent pas être soumis au contrôle des laboratoires pour y être expérimentés ; parce que nous avons de nombreux chromosomes et encore plus de caractères héréditaires que la mouche du vinaigre ; parce que les traits de caractère humains qui nous intéressent le plus – le génie créatif, l'intelligence, la valeur morale – sont extrêmement complexes, impliquent l'interaction de nombreux gènes et subissent l'influence du milieu ; pour toutes ces raisons donc, les généticiens ne peuvent pas venir à bout de la génétique humaine avec la même assurance que celle dont ils font preuve dans la génétique de la mouche du vinaigre.

L'eugénisme demeure donc un rêve, flou en raison du manque de connaissances, et pervers car il peut être facilement utilisé par les racistes et les fanatiques.

LA GÉNÉTIQUE MOLÉCULAIRE

Comment un gène produit-il la caractéristique physique dont il est responsable dans un être ? Par quel mécanisme donne-t-il naissance à la couleur jaune des graines de pois, ou aux ailes frisées de la mouche du vinaigre, ou aux yeux bleus des êtres humains ?

Les biologistes sont maintenant certains que les gènes exercent leurs effets au moyen des enzymes. Un des exemples les plus évidents est la couleur des yeux, des cheveux et de la peau. La couleur (bleue ou brune,

jaune ou noire, ou les teintes intermédiaires) est déterminée par la quantité d'un pigment, la *mélanine* (du mot grec pour « noir »), qui est présent dans l'iris de l'œil, les cheveux, la peau. La mélanine est synthétisée à partir d'un acide aminé, la tyrosine, au cours d'un certain nombre d'étapes qui ont été déterminées, qui sont effectuées par des enzymes, et la quantité de mélanine formée dépend de celle des enzymes. Par exemple, l'une de ces enzymes, la tyrosinase, catalyse les deux premières étapes. Un gène déterminé contrôle la production de la tyrosinase par la cellule, et contrôle ainsi la couleur de la peau, des yeux et des cheveux. Comme le gène est transmis de génération en génération, les enfants auront naturellement une couleur semblable à celle de leurs parents. S'il se trouve une mutation qui produit un gène défectueux qui ne peut pas former de tyrosinase, la mélanine sera absente, et l'individu sera *albinos*. L'absence d'une seule enzyme, et donc la déficience d'un seul gène, suffit à causer un changement majeur de l'aspect d'un individu.

Si l'on part du principe que les caractéristiques d'un individu dépendent de son contenu en enzymes, qui lui-même est contrôlé par les gènes, la question qui se pose est la suivante : comment les gènes fonctionnent-ils ? Malheureusement, même la mouche du vinaigre est un organisme trop compliqué pour y étudier ce sujet en détail. Mais, en 1941, les biologistes américains George Wells Beadle et Edward Lawrie Tatum se servirent d'un organisme simple qu'ils trouvèrent admirablement adapté à ce but : la moisissure du pain (*Neurospora crassa*).

Neurospora n'est pas très exigeante quant à sa nourriture. Elle pousse très bien sur du sucre additionné de composés qui lui fournissent de l'azote, du soufre et quelques éléments minéraux. A part le sucre, la seule substance organique qui doit lui être fournie est une vitamine, *la biotine*.

A un certain stade de son cycle, la moisissure produit huit spores, toutes de composition génétique identique. Chaque spore contient sept chromosomes ; comme dans la cellule germinale des organismes supérieurs, leurs chromosomes sont uniques et non pas appariés. En conséquence, si l'un des chromosomes est modifié, l'effet de cette modification peut être observé, parce qu'il n'y a pas de partenaire chromosomique normal pour masquer cet effet. Beadle et Tatum créèrent des mutations chez Neurospora en exposant la moisissure aux rayons X, et en suivirent ensuite les effets spécifiques sur le comportement des spores.

Si la moisissure a reçu une dose de radiations et que les spores peuvent encore se développer sur le milieu nutritionnel habituel, elle ne subit pas de mutation, tout au moins en ce qui concerne les exigences nutritionnelles de l'organisme pour sa croissance. Si les spores ne poussent pas sur le milieu habituel, l'expérimentateur détermine si elles sont vivantes ou mortes en les cultivant avec un milieu complet, contenant toutes les vitamines, acides aminés et autres ingrédients dont elles pourraient éventuellement avoir besoin. Si les spores poussent sur le milieu, on en conclut que les rayons X ont produit une mutation qui modifie les besoins nutritionnels de Neurospora. Apparemment, elle a besoin maintenant d'un nouvel ingrédient dans son régime. Pour savoir lequel, les chercheurs ont procédé par élimination ; ils ont fait pousser les spores sur des milieux complets d'où chaque fois ils retiraient un ingrédient différent. Ils ont ainsi supprimé soit tous les acides aminés, soit les différentes vitamines, soit un ou deux acides aminés ou une ou deux vitamines. Ils ont finalement réduit le nombre d'ingrédients possibles jusqu'à pouvoir

identifier exactement ce dont la spore a besoin dans le milieu pour compenser la mutation.

Il semble parfois que la spore mutée ait besoin d'arginine. La souche *sauvage* normale fabrique son arginine à partir du sucre et de sels d'ammonium. A cause de la modification génétique, elle ne peut plus synthétiser d'arginine ; et à moins que l'acide aminé ne lui soit fourni dans le milieu, elle ne peut pas fabriquer de protéines et donc croître.

Pour rendre compte d'une telle situation, il suffit de supposer que les rayons X ont détruit le gène responsable de la formation d'une enzyme nécessaire à la synthèse de l'arginine. En l'absence du gène normal, *Neurospora* ne peut plus fabriquer la bonne enzyme. Pas d'enzyme, pas d'arginine.

Beadle et ses collaborateurs utilisèrent ce type d'information pour étudier la relation entre les gènes et la chimie du métabolisme. Ils ont pu démontrer, par exemple, que plus d'un gène est impliqué dans la synthèse de l'arginine. Pour simplifier, disons qu'il y a deux gènes, A et B, responsables de la formation de deux enzymes différentes, toutes deux nécessaires à la synthèse de l'arginine. Une mutation soit dans A, soit dans B, empêchera *Neurospora* de produire l'acide aminé. Supposons que nous irradiions deux groupes de *Neurospora* et que nous obtenions une souche incapable de synthétiser l'arginine dans aucun des deux. Si nous avons de la chance, un mutant aura un gène A défectueux et un gène B normal ; et l'autre un gène A normal et un gène B défectueux. Pour voir ce qui s'est passé, croisons les deux mutants au stade sexuel de leur cycle. Si les deux souches diffèrent vraiment de cette façon, la recombinaison des chromosomes peut produire des spores où les deux gènes A et B sont normaux. En d'autres termes, on peut obtenir à partir de deux mutants incapables de pousser sans arginine des descendants qui le *peuvent*. Bien sûr, c'est exactement ce qui s'est passé quand les expériences ont été réalisées.

On peut étudier, encore plus en détail, le métabolisme de *Neurospora*. Par exemple, on connaît trois mutants incapables de synthétiser de l'arginine à partir d'un milieu normal. L'un ne peut croître que si on lui fournit l'arginine. Le second ne peut pousser que si on lui donne de l'arginine ou un composé voisin, *la citrulline*. Le troisième pousse sur un milieu contenant soit de l'arginine, soit de la citrulline, soit encore un autre composé similaire, *l'ornithine*.

Quelles conclusions peut-on tirer de tout cela ? On peut deviner que ces trois substances sont les étapes d'une séquence dont l'arginine est le produit final, chacune ayant besoin d'une enzyme. Tout d'abord, l'ornithine est formée à partir d'un composé plus simple à l'aide d'une enzyme ; ensuite, une autre enzyme convertit l'ornithine en citrulline ; enfin, une troisième enzyme convertit la citrulline en arginine. Un mutant de *Neurospora* auquel il manque l'enzyme pour faire l'ornithine, mais possède les autres enzymes, peut pousser si on lui fournit de l'ornithine, car à partir de là, la spore peut faire la citrulline, puis l'arginine essentielle. Naturellement, le mutant peut aussi pousser sur de la citrulline, dont il peut faire de l'arginine. On peut raisonner de la même manière en disant que la deuxième souche mutante n'a pas l'enzyme nécessaire à la conversion de l'ornithine en citrulline. On doit donc lui fournir de la citrulline, dont elle fait de l'arginine, ou bien l'arginine elle-même. Finalement, on peut conclure que le mutant qui ne croît que sur de

l'arginine a perdu l'enzyme (et le gène) responsable de la conversion de la citrulline en arginine.

En analysant le comportement des diverses souches mutantes, qu'ils ont isolées, Beadle et ses collaborateurs ont fondé la *génétique moléculaire*. Ils ont déterminé les voies utilisées par les organismes pour synthétiser un certain nombre de composés importants. Beadle a proposé la *théorie du un-gène-une-enzyme*, selon laquelle chaque gène gouverne la formation d'une seule enzyme ; une hypothèse qui est maintenant généralement acceptée par les généticiens. Beadle et Tatum reçurent le prix Nobel de médecine et de physiologie en 1958 pour ce travail de pionniers.

LES HÉMOGLOBINES ANORMALES

Les découvertes de Beadle ont convaincu les biochimistes que chez l'homme aussi, les modifications sont dues à des mutations des gènes qui les gouvernent. Une preuve se présenta inopinément, en relation avec une maladie, *l'anémie falciforme*, l'une des quelque 1 600 maladies génétiques humaines connues à l'heure actuelle.

Cette maladie a été décrite pour la première fois en 1910 par un médecin de Chicago, James Bryan Herrick. Alors qu'il examinait au microscope un échantillon de sang d'adolescent noir, Herrick observa que ses globules rouges, qui sont normalement ronds, avaient la forme un peu courbée d'une faucille. D'autres médecins remarquèrent que ce même phénomène se produisait presque toujours chez des patients noirs. Les chercheurs conclurent finalement que l'anémie falciforme est une maladie héréditaire. Sa transmission suit les lois de Mendel ; apparemment, il y a un gène de l'anémie qui, lorsqu'il est hérité des deux parents, provoque cette déformation des globules. Ces globules sont incapables de transporter l'oxygène correctement, et ont un temps de vie exceptionnellement court, de sorte qu'il y a un déficit de globules rouges dans le sang. Ceux qui héritent de deux copies du gène meurent souvent de cette maladie dès l'enfance. D'autre part, quand une personne n'a qu'une seule copie du gène de l'anémie falciforme, celle d'un seul de ses parents, la maladie n'apparaît pas. La forme en faucille des globules rouges ne se produit que lorsque la personne est privée d'oxygène à un degré anormal, comme en altitude. On considère alors que la personne a *le trait falciforme*, mais pas la maladie.

On a trouvé que 9 % environ de la population noire américaine a le trait falciforme, et que 0,25 % a la maladie. Dans certaines localités d'Afrique Centrale, jusqu'à 25 % de la population est porteuse de ce trait. Le gène de l'anémie falciforme serait apparu à la suite d'une mutation quelque part en Afrique et se serait propagé depuis chez les individus d'ascendance africaine. Si la maladie est fatale, comment se fait-il que ce gène défectueux ne se soit pas éteint ? Des études faites en Afrique dans les années cinquante ont apporté la réponse. Il semble que les sujets porteurs du trait de l'anémie falciforme soient plus résistants à la malaria que les individus normaux. Les globules falciformes sont inhospitaliers au parasite de la malaria. On estime que, dans les régions infestées par la malaria, les enfants possédant ce trait ont 25 % de chances de plus de survivre jusqu'à l'âge de la procréation que les enfants non porteurs. Donc, posséder une seule copie du gène de l'anémie falciforme (mais pas les deux copies, qui donnent la maladie) confère un avantage. Les deux tendances

opposées – promotion du gène défectueux par l'effet protecteur d'une seule copie, et élimination du gène par son effet fatal lorsqu'il est présent en deux copies – tend à créer un équilibre du gène dans la population.

Dans les régions où la malaria ne constitue pas un problème grave, le gène a en effet tendance à s'éteindre. En Amérique, dans la population noire, l'incidence du gène de l'anémie falciforme a peut-être été au départ de 25 %. Même en tenant compte d'une diminution à 15 % à cause du métissage, l'incidence actuelle de 9 % montre que le gène disparaît. En toute probabilité, il continuera à le faire. Si l'Afrique est libérée de la malaria, le gène s'y éteindra aussi.

La signification biochimique du gène de l'anémie falciforme passa brutalement au premier plan en 1949 ; Linus Pauling et ses collaborateurs de l'Institut de technologie de Californie (où Beadle travaillait aussi) démontrèrent en effet que le gène de l'anémie falciforme affecte l'hémoglobine des globules rouges : les personnes ayant deux copies du gène sont incapables de produire de l'hémoglobine normale. Pauling l'a prouvé en se servant d'une technique, *l'électrophorèse,* qui sépare les protéines au moyen du courant électrique selon leur différence de charge électrique nette (cette technique a été mise au point par le chimiste suédois Arne Wilhelm Kaurin Tiselius, qui reçut le prix Nobel de chimie en 1948 pour cette précieuse contribution). Pauling, à l'aide de l'analyse électrophorétique, découvrit que les patients atteints d'anémie falciforme ont une hémoglobine anormale (*l'hémoglobine S*), qui peut être séparée de l'hémoglobine normale. Cette dernière est appelée *hémoglobine A* (pour « adulte »), ce qui la distingue des hémoglobines fœtales, les *hémoglobines F.*

Depuis 1949, les biochimistes ont découvert des hémoglobines anormales autres que celle de l'anémie falciforme, auxquelles ils ont donné les lettres de C à M. Il semble que le gène responsable de la synthèse de l'hémoglobine ait été muté, créant de nombreux allèles différents ; chacun donne naissance à une hémoglobine qui a des capacités moindres de transport de l'oxygène dans les conditions normales, mais peut être utile dans certaines conditions inhabituelles. Tout comme l'hémoglobine S, présente en une seule copie, confère une meilleure résistance à la malaria, l'hémoglobine C, en une seule copie, permet à l'organisme de vivre dans des conditions où le fer n'est présent qu'en quantité marginale.

Étant donné que les diverses hémoglobines anormales diffèrent par leur charge électrique, on peut supposer qu'elles diffèrent aussi dans l'arrangement des acides aminés de la chaîne peptidique, car le contenu en acides aminés est responsable de la charge de la molécule. Ces différences doivent être très réduites, puisque les hémoglobines anormales fonctionnent toutes plus ou moins comme l'hémoglobine. L'espoir de localiser cette différence dans une molécule énorme contenant quelque six cents acides aminés semblait donc très faible. Vernon Martin Ingram, un biochimiste germano-américain, et ses collaborateurs se sont attaqués malgré tout au problème.

Ils ont coupé les hémoglobines A, S et C avec une enzyme qui dégrade les protéines en peptides de différentes tailles. Ils ont séparé les peptides de chaque hémoglobine par une *électrophorèse sur papier,* c'est-à-dire qu'ils ont utilisé le courant, au lieu d'une solution, pour transporter les peptides le long d'un morceau de papier humide (chromatographie sur papier électrifié). En répétant l'expérience avec les trois hémoglobines, les chercheurs ont trouvé qu'elles ne différaient que par un seul peptide que l'on retrouvait chaque fois à un emplacement différent.

Ils ont dégradé et analysé ce peptide. Ils ont appris qu'il était composé de neuf acides aminés, et que l'organisation de ces acides aminés était la même dans les trois hémoglobines, à une exception près. Les arrangements respectifs sont :

Hémoglobine A : His-Val-Leu-Leu-Thr-Pro-Glu-Glu-Lys
Hémoglobine S : His-Val-Leu-Leu-Thr-Pro-Val-Glu-Lys
Hémoglobine C : His-Val-Leu-Leu-Thr-Pro-Lys-Glu-Lys

La seule différence entre les trois hémoglobines réside dans l'acide aminé situé à la septième position du peptide ; il s'agit de l'acide glutamique pour l'hémoglobine A, la valine pour l'hémoglobine S, et la lysine pour l'hémoglobine C. Étant donné que l'acide glutamique fournit une charge négative, la lysine une charge positive et la valine pas de charge du tout, on comprend pourquoi les trois protéines ont un comportement différent en électrophorèse : leur distribution de charge est différente.

Mais pourquoi un changement aussi insignifiant dans une molécule conduit-il à un changement aussi important au niveau du globule rouge ? Le globule rouge normal est constitué pour un tiers d'hémoglobine A. Les molécules d'hémoglobine A sont si serrées dans le globule rouge qu'elles ont à peine la place de bouger. En bref, elles sont à la limite de la précipitation. Le caractère de la charge de la protéine, entre autres, détermine si une protéine précipitera ou non. Si toutes les protéines ont la même charge, elles se repoussent les unes les autres et empêchent la précipitation. Plus la charge est grande (c'est-à-dire la *répulsion*), plus elles ont de chances de rester en solution. Pour l'hémoglobine S, la répulsion intermoléculaire est peut-être légèrement plus faible que pour l'hémoglobine A, et l'hémoglobine est proportionnellement moins soluble et plus à même de précipiter. Quand un gène de l'anémie falciforme est couplé à un gène normal, ce dernier forme sans doute assez d'hémoglobine A pour maintenir l'hémoglobine S en solution, mais il s'en faut de peu. Mais quand les deux gènes sont mutants, ils ne produisent que de l'hémoglobine S, qui ne peut rester en solution. Elle précipite, formant des cristaux qui causent une distorsion et affaiblissent le globule rouge.

Cette théorie expliquerait pourquoi le changement d'un seul acide aminé dans chacune des demi-molécules composées de près de six cents acides aminés est suffisant pour produire une maladie grave et la quasi-certitude d'une mort prématurée.

LES ANOMALIES MÉTABOLIQUES

L'albinisme et l'anémie falciforme ne sont pas les seules maladies dues à l'absence d'une seule enzyme ou à la mutation d'un seul gène. On connaît la *phénylcétonurie,* une anomalie héréditaire du métabolisme qui provoque souvent une arriération mentale, et qui résulte de l'absence de l'enzyme nécessaire à la conversion de la phénylalanine en tyrosine. Il y a la *galactosémie,* un trouble causant la cataracte de l'œil et des dommages au foie et au cerveau, qui est dû à l'absence de l'enzyme qui convertit le galactose phosphate en glucose phosphate. Il existe des anomalies qui consistent en l'absence de l'une ou l'autre des enzymes qui contrôlent la dégradation du *glycogène* (une sorte d'amidon) en glucose, ce qui donne une accumulation anormale de glycogène dans le foie et les muscles, et

entraîne souvent une mort précoce. Ce sont tous des exemples de *défauts congénitaux du métabolisme,* d'une incapacité innée à former des enzymes existant chez l'homme normal, et qui sont plus ou moins vitales. Ce concept fut introduit en médecine en 1908 par Archibald Edward Garrod, un médecin anglais, mais il resta ignoré pendant toute une génération jusqu'à ce que vers les années trente, le généticien anglais John Burdon Sanderson Haldane porte à nouveau le sujet à l'attention des chercheurs.

Ce genre de troubles sont généralement contrôlés par des allèles récessifs du gène qui produit l'enzyme en question. Quand un seul des deux gènes de la paire est défectueux, le gène normal fonctionne et l'individu est en général capable de mener une vie normale (comme c'est le cas du porteur du trait de l'anémie falciforme). Les problèmes ne surviennent en général que lorsque les deux parents se trouvent avoir le même gène anormal et la malchance supplémentaire de combiner les deux dans un œuf fécondé. Leur enfant est alors victime d'une sorte de roulette russe. Nous portons probablement tous notre fardeau de gènes anormaux, défectueux et même dangereux, habituellement masqués par des gènes normaux. On peut comprendre pourquoi les chercheurs en génétique humaine sont si préoccupés des radiations et de tout ce qui peut augmenter le taux de mutations.

Les acides nucléiques

Ce qui est surprenant, dans l'hérédité, ce ne sont pas ces anomalies spectaculaires, mais relativement rares ; c'est plutôt le fait que, à tout prendre, l'hérédité fonctionne si fidèlement vis-à-vis de l'espèce. Génération après génération, millénaire après millénaire, les gènes se reproduisent sous la même forme exactement, et génèrent exactement les mêmes enzymes, avec seulement, occasionnellement, une variation accidentelle du schéma directeur. L'introduction d'un seul acide aminé erroné dans une grande protéine se produit rarement. Comment les gènes s'arrangent-ils pour faire ainsi des copies aussi fidèles d'eux-mêmes à l'infini ?

La réponse doit se trouver dans la structure des longs filaments de gènes que nous appelons chromosomes. La plus grande partie du chromosome, environ la moitié de sa masse, est constituée de protéines. Ce n'est pas une surprise. A mesure que le XXe siècle avançait, les biochimistes s'attendaient à ce que les protéines soient partie intégrante de toute fonction complexe de l'organisme. Les protéines semblaient être les molécules complexes de l'organisme, les seules suffisamment complexes pour répondre à toute la souplesse et la sensibilité de la vie.

Pourtant, la majeure partie des protéines du chromosome appartiennent à une catégorie appelée *histones,* dont les molécules sont plutôt petites pour des protéines et (pire encore) constituées d'un mélange très simple d'acides aminés. Elles ne semblaient pas assez complexes pour être responsables de la complexité et de la finesse de la génétique. Il y a, bien sûr, des protéines non histones, qui sont constituées de molécules beaucoup plus grandes et complexes, mais elles ne représentent qu'une faible proportion de l'ensemble.

Les biochimistes étaient cependant en panne avec les protéines. Les mécanismes de l'hérédité ne pouvaient impliquer rien d'autre que les

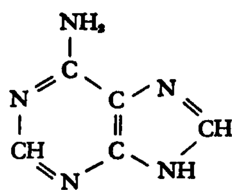
protéines. A peu près la moitié du chromosome était constituée d'une substance qui n'avait rien à voir avec les protéines, mais il ne semblait pas possible d'envisager autre chose que les protéines. Nous devons tout de même aller voir du côté de ce constituant non protéique du chromosme.

LA STRUCTURE GÉNÉRALE

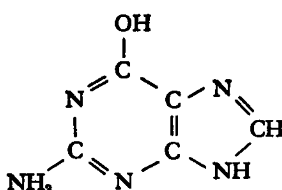
En 1869, Friedrich Miescher, un biochimiste suisse, alors qu'il dégradait les protéines des cellules avec de la pepsine, remarqua que la pepsine ne cassait pas le noyau cellulaire. Le noyau se rétrécissait un peu, mais il restait intact. Par l'analyse chimique, Miescher trouva alors que le noyau de la cellule consistait essentiellement en une substance contenant du phosphore, dont les propriétés ne ressemblaient pas du tout à celles des protéines. Il l'appela *nucléine.* Elle fut rebaptisée *acide nucléique* vingt ans plus tard, quand on découvrit qu'elle était fortement acide.

Miescher se consacra entièrement à l'étude de ce nouveau matériel et découvrit ainsi que le spermatozoïde (presque entièrement constitué de matériel nucléaire) est particulièrement riche en acides nucléiques. Sur ces entrefaites, le chimiste allemand Felix Hoppe-Seyler, dans le laboratoire duquel Miescher avait fait ses premières découvertes, et qui avait personnellement confirmé le travail du jeune chercheur avant de le laisser publier, isola l'acide nucléique de la levure. Ses propriétés semblaient différentes de celles du matériel de Miescher, qui fut appelé *acide nucléique du thymus* (parce qu'il pouvait être obtenu aisément du thymus des animaux) ; celui de Hoppe-Seyler fut naturellement appelé *acide nucléique de la levure.* Comme l'acide nucléique du thymus n'a été obtenu au départ que de cellules animales et celui de la levure de cellules de plantes, on a pensé pendant un certain temps que cela représentait une distinction chimique générale entre les animaux et les végétaux.

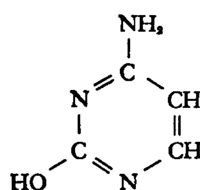
Le biochimiste allemand Albrecht Kossel, un autre élève de Hoppe-Seyler, fut le premier à faire une étude systématique de la structure des acides nucléiques. A la suite d'une hydrolyse minutieuse, il en isola une série de composés azotés, qu'il nomma *adénine, guanine, cytosine et thymine.* On sait maintenant que leurs formules sont les suivantes :



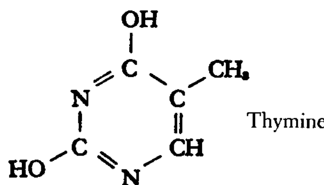
Adénine



Guanine



Cytosine

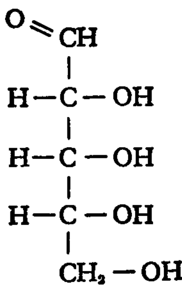


Thymine

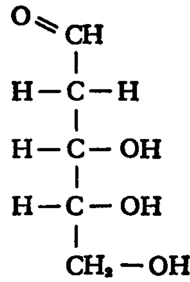
L'ensemble des deux noyaux des deux premiers composés constitue le *noyau purique,* et l'unique noyau des deux derniers le *noyau pyrimidique.* On se réfère donc à la guanine et à l'adénine en tant que *purines* et à la cytosine et à la thymine en tant que *pyrimidines.*

Pour ces recherches qui furent le départ d'une série de découvertes fructueuses, Kossel reçut le prix Nobel de médecine et de physiologie en 1910.

En 1911, le biochimiste américain d'origine russe Phœbus Aaron Theodore Levene, un élève de Kossel, franchit un pas de plus. Kossel avait découvert en 1891 que les acides nucléiques contiennent des sucres, et Levene remarqua que les acides nucléiques contiennent un sucre à cinq carbones. C'était à cette époque une découverte insolite : les sucres les plus connus, tel le glucose, contiennent six carbones. Levene poursuivit ses recherches en montrant que les deux variétés d'acides nucléiques diffèrent par la nature de leur sucre à cinq carbones. L'acide nucléique de la levure contient du *ribose,* tandis que l'acide nucléique du thymus contient un sucre qui ressemble beaucoup au ribose, sauf qu'il lui manque un atome d'oxygène ; il fut donc appelé *désoxyribose.* Leurs formules sont :



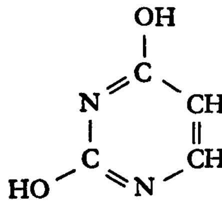
Ribose



Déoxyribose

En conséquence, les deux acides nucléiques furent appelés *acide ribonucléique* (ARN) et *désoxyribonucléique* (ADN).

En dehors de leurs deux sucres, les deux acides nucléiques diffèrent aussi par l'une des pyrimidines. L'ARN a de l'*uracile* à la place de la thymine. L'uracile ressemble à la thymine, comme le montre la formule :



Uracile

Dès 1934, Levene montra que les acides nucléiques peuvent être cassés en fragments qui contiennent une purine ou une pyrimidine, et soit du ribose soit du désoxyribose, plus un groupement phosphate. Cette association constitue un *nucléotide.* Levene suggéra que les molécules d'acide nucléique étaient constituées de nucléotides comme les protéines sont constituées d'acides aminés. D'après ses études quantitatives, la

molécule n'était composée que de quatre unités de nucléotides, l'une contenant l'adénine, une la guanine, une la cytosine et une, soit la thymine (dans l'ADN), soit l'uracile (dans l'ARN).

Cette proposition semblait logique. On pensait que le matériel des chromosomes, entre autres, était constitué de *nucléoprotéines,* qui à leur tour étaient constituées de grandes molécules de protéines auxquelles étaient attachés un ou plusieurs de ces groupements *tétranucléotides,* qui servaient à une fonction inconnue mais subsidiaire.

Il s'avéra en fait que ce que Levene avait isolé n'était pas des molécules d'acides nucléiques, mais des fragments ; au milieu des années cinquante, les biochimistes trouvèrent que le poids moléculaire des acides nucléiques pouvait dépasser les six millions. Les acides nucléiques ont donc un poids moléculaire aussi élevé, si ce n'est plus élevé que les protéines.

La manière exacte dont les nucléotides sont enchaînés a été confirmée par le biochimiste anglais Alexander Robertus Todd, qui a construit une série de nucléotides à partir de fragments plus simples et les a liés entre eux avec soin, dans des conditions ne permettant qu'à un seul type de liaison de se faire. Cela lui a valu le prix Nobel de chimie en 1957.

En conséquence, on peut envisager une structure générale des acides nucléiques ressemblant en quelque sorte à celle des protéines. La molécule protéique est constituée d'une colonne vertébrale polypeptidique d'où dépassent les chaînes latérales des acides aminés. Dans les acides nucléiques, le sucre d'un nucléotide est lié au sucre du suivant par l'intermédiaire du groupement phosphate qui est attaché aux deux. Une *ossature de sucre-phosphate* parcourt la molécule sur sa longueur et il en émerge les purines et les pyrimidines, une par nucléotide.

Les nucléoprotéines sont donc constituées de deux parties qui sont de grandes *macromolécules.* La fonction de la partie acide nucléique devint une question de premier ordre.

L'ADN

En utilisant les techniques de coloration des cellules, les chercheurs localisèrent les acides nucléiques dans la cellule. Le chimiste allemand Robert Feulgen utilisa un colorant rouge qui colorait l'ADN, mais pas l'ARN, et montra ainsi que l'ADN est localisé dans le noyau cellulaire, plus précisément dans les chromosomes. Il le détecta dans les cellules animales et les cellules végétales. De plus, en colorant l'ARN, il démontra que cet acide nucléique est aussi présent dans les cellules animales et les cellules végétales. En bref, les acides nucléiques sont un matériel universel présent dans toutes les cellules.

Le biochimiste suédois Torbjörn Caspersson étudia le sujet plus en détail en éliminant un des acides nucléiques (au moyen d'une enzyme qui le réduit en fragments solubles qui peuvent être entraînés par lavage hors de la cellule) et en se concentrant sur le deuxième. Il photographia les cellules en lumière ultraviolette (les acides nucléiques absorbant les ultraviolets bien davantage que n'importe quel autre matériel cellulaire), et localisa l'ADN ou l'ARN (celui qu'il avait laissé dans la cellule). Avec cette technique, l'ADN n'était visible que dans les chromosomes, et l'ARN principalement dans certaines particules cytoplasmiques. On trouvait aussi de l'ARN dans le *nucléole,* une structure du noyau. En 1948, Alfred Ezra

Mirsky, un biochimiste du Rockefeller Institute, montra la présence de faibles quantités d'ARN dans les chromosomes, tandis que Ruth Sager affirma qu'il y a de l'ADN dans le cytoplasme, notamment dans les chloroplastes des plantes. En 1966, l'ADN était aussi localisé dans les mitochondries.

Les photos de Caspersson révélèrent que dans le chromosome, l'ADN se présente sous forme de bandes. L'ADN pourrait-il être le support des gènes, dont l'existence était restée plutôt vague jusqu'alors ?

Au cours des années quarante, les biochimistes poursuivirent cette piste avec un enthousiasme grandissant. Il était pour eux particulièrement significatif que la quantité d'ADN dans les cellules d'un organisme soit strictement constante, à l'exception des cellules germinales qui n'en possèdent que la moitié, comme on s'y attendait, puisqu'elles ont moitié moins de chromosomes que les cellules normales. Les quantités d'ARN et de protéines peuvent varier largement, mais la quantité d'ADN reste fixe. Cela permettait d'établir une relation étroite entre les gènes et l'ADN.

Le rôle des protéines comme supports des gènes commençait à vaciller sous les coups de boutoir des acides nucléiques ; c'est alors que des observations exceptionnelles furent rendues publiques, qui affirmaient que l'ADN *était* le gène.

Les bactériologistes étudiaient depuis longtemps deux souches différentes de pneumocoques cultivées en laboratoire : l'une présentait une paroi lisse faite de glucides complexes ; l'autre n'avait pas de paroi et présentait une apparence rugueuse. Il semblait qu'une enzyme nécessaire à la confection de la paroi soit absente dans la souche rugueuse. Néanmoins, Fred Griffith, un bactériologiste anglais, découvrit que s'il mélangeait des bactéries mortes de la souche lisse avec des bactéries rugueuses vivantes et s'il injectait le mélange dans une souris, les tissus de la souris infectée contenaient des bactéries lisses vivantes ! Comment cela s'était-il produit ? Les bactéries mortes n'avaient sûrement pas ressuscité. Quelque chose avait dû transformer les bactéries rugueuses de telle sorte qu'elles soient ensuite capables de synthétiser une paroi lisse. De quoi s'agissait-il ? C'était bien sûr un facteur fourni par le pneumocoque mort de la souche lisse.

En 1944, trois biochimistes américains, Oswald Theodore Avery, Colin Munro Macleod et Maclyn McCarty, identifièrent le principe transformant : c'était de l'ADN. En isolant de l'ADN pur de la souche lisse et en le donnant au pneumocoque rugueux, ils arrivaient à transformer la souche rugueuse en souche lisse.

Les chercheurs isolèrent alors d'autres principes transformants, à partir d'autres bactéries, touchant d'autres propriétés, et chaque fois le principe s'avéra être une variété d'ADN. La seule conclusion plausible était que l'ADN peut agir comme un gène. En fait, différents projets de recherche, en particulier sur les virus (voir chapitre 14), montrèrent que les protéines associées à l'ADN sont presque superflues d'un point de vue génétique ; l'ADN peut produire par lui-même les effets génétiques, soit sous la forme du chromosome, soit – dans les cas d'hérédité non chromosomique – sous la forme de corpuscules cytoplasmiques tels que les chloroplastes et les mitochondries.

LA DOUBLE HÉLICE

Si l'ADN est la clef de l'hérédité, il doit avoir une structure complexe, car il lui faut porter un schéma élaboré, un code d'instructions (le *code*

génétique) pour la synthèse des enzymes. S'il n'est constitué que de quatre sortes de nucléotides, ils ne peuvent être enchaînés en un arrangement régulier tel que 1, 2, 3, 4, 1, 2, 3, 4, 1, 2, 3, 4... Une molécule de ce genre serait beaucoup trop simple pour être la matrice des enzymes. En fait, le biochimiste américain Erwin Chargaff et ses collaborateurs prouvèrent définitivement, en 1948, que la composition des acides nucléiques était bien plus complexe qu'on ne l'avait supposé. Leur analyse montra que les différentes purines et pyrimidines ne sont pas présentes en quantités égales, mais que leur proportion varie dans les différents acides nucléiques.

Tout semblait démontrer que les quatre purines et pyrimidines sont distribuées au hasard le long de l'ADN comme les acides aminés dans les protéines ; néanmoins, une certaine régularité semblait exister. Dans toutes les molécules d'ADN, le nombre total de purines semble toujours égal au nombre total de pyrimidines. De plus, le nombre d'adénines (une purine) est toujours égal au nombre de thymines (une pyrimidine), tandis que le nombre de guanines (l'autre purine) est toujours égal au nombre de cytosines (l'autre pyrimidine).

En représentant l'adénine par A, la guanine par G, la thymine par T et la cytosine par C, les purines seront alors A + G et les pyrimidines T + C. Les découvertes concernant toutes les molécules d'ADN peuvent être résumées ainsi :

$$A = T$$
$$G = C$$
$$A + G = T + C$$

D'autres symétries sont aussi apparues. Dès 1938, Astbury avait signalé que les acides nucléiques donnaient un diagramme de diffraction des rayons X, un signe de l'existence de symétries structurales dans la molécule. Maurice Hugh Frederick Wilkins, un biochimiste anglais né en Nouvelle-Zélande, calcula que ces répétitions se retrouvaient à des intervalles bien plus grands que la distance d'un nucléotide à l'autre. Selon une conclusion logique, la molécule d'ADN prend la forme d'une hélice, l'enroulement de l'hélice formant l'unité répétitive mise en évidence par les rayons X. Cette hypothèse était tout à fait plausible car Linus Pauling démontrait au même moment la structure en hélice de certaines molécules de protéines.

Les conclusions de Wilkins étaient largement basées sur les travaux de son associée, Rosalind Elsie Franklin, dont le rôle dans ces études fut constamment minimisé, principalement en raison de l'attitude mysogine des autorités scientifiques britanniques.

En 1953, le physicien anglais Francis Harry Compton Crick et son collaborateur, l'Américain James Dewey Watson, biochimiste et ancien enfant prodige, rassemblèrent tous les éléments d'information disponibles, utilisant une photographie clé prise par Franklin, apparemment sans sa permission, et arrivèrent à un modèle révolutionnaire de la molécule d'acide nucléique. Ce modèle ne la représentait pas sous la forme d'une simple hélice (là était le point essentiel), mais sous la forme d'une double hélice : deux ossatures de sucre-phosphate tournant le long d'un axe vertical comme la double spirale d'un escalier (figure 13.6). Branchées sur l'ossature de sucre-phosphate, les purines et les pyrimidines se dirigent vers l'intérieur, les unes vers les autres, se joignant pour constituer les marches de cet escalier en colimaçon.

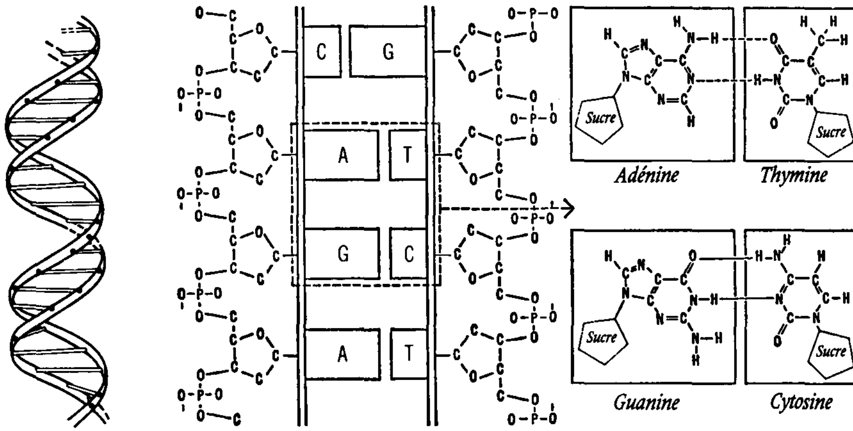


Figure 13.6. Le modèle de l'acide désoxyribonucléique. Le dessin *à gauche* représente la double hélice ; *au centre,* agrandissement d'une partie de la molécule (les atomes d'hydrogène sont omis) ; *à droite,* le détail de l'appariement des nucléotides.

Comment les purines et les pyrimidines arrivent-elles à s'ordonner le long de ces chaînes parallèles ? Pour constituer un bon assemblage uniforme, une purine à deux noyaux d'un côté doit toujours faire face à une pyrimidine à un noyau de l'autre. Deux pyrimidines ne suffiraient pas à couvrir l'espace ; tandis que deux purines seraient trop entassées. De plus, l'adénine d'une chaîne doit toujours faire face à une thymine de l'autre chaîne, et une guanine être en face d'une cytosine. On peut ainsi expliquer que $A = T$, $G = C$ et $A + T = G + C$.

Le *modèle de Watson et Crick* de la structure de l'ADN s'est révélé extrêmement fructueux ; et Wilkins, Crick et Watson partagèrent le prix Nobel de médecine et de physiologie en 1962. Franklin étant morte en 1958, la question de sa contribution ne fut pas soulevée.

Le modèle de Watson et Crick permet, par exemple, d'expliquer comment un chromosome peut se dupliquer pendant la division cellulaire. Si l'on considère le chromosome comme une chaîne de molécules d'ADN, les molécules peuvent d'abord se diviser en ouvrant la double hélice en deux, les deux chaînes se détachant (en se déroulant) l'une de l'autre (un peu à la manière d'une fermeture éclair). Ceci peut avoir lieu parce que les pyrimidines et les purines, qui sont face à face, sont maintenues par des liaisons hydrogènes suffisamment faibles pour être facilement rompues. Chaque chaîne est une demi-molécule qui permet la synthèse de sa partie complémentaire manquante. Là où une thymine est présente, une adénine se lie ; là où se trouve une cytosine, une guanine se lie, etc. Tous les matériaux de base et les enzymes nécessaires sont présents dans la cellule. La demi-molécule joue simplement le rôle de *matrice* ou de moule pour associer les unités dans le bon ordre. Les unités se retrouvent finalement à la place appropriée et y restent parce que c'est l'arrangement le plus stable.

En résumé, chaque demi-molécule guide la formation de son propre complément, qui s'y lie par liaisons hydrogènes. La double hélice complète

est ainsi reconstituée, et les deux demi-molécules provenant de la molécule originale forment deux molécules là où il n'en existait qu'une auparavant. Un tel processus, appliqué à tout l'ADN du chromosome, crée deux chromosomes qui sont exactement semblables et des copies parfaites du chromosome père original. Occasionnellement, un facteur défavorable intervient : l'impact d'une particule subatomique, l'énergie d'une radiation ou l'intervention de certains produits chimiques peuvent introduire une imperfection à un endroit quelconque dans le nouveau chromosome. Le résultat est une mutation.

Les preuves en faveur de ce mécanisme de réplication se sont accumulées. Des études réalisées avec l'azote lourd comme traceur pour marquer les chromosomes et suivre le devenir du matériel marqué pendant la division cellulaire, étayent cette théorie. De plus, on a pu ainsi identifier certaines enzymes très importantes dans le processus de la réplication.

En 1955, le biochimiste hispano-américain Severo Ochoa isola d'une bactérie *(Aztobacter vinelandii)* une enzyme capable de catalyser la formation d'ARN à partir de nucléotides. En 1956, un ancien élève d'Ochoa, Arthur Kornberg, isola une autre enzyme (de la bactérie *Escherichia coli*), qui catalyse la formation d'ADN à partir de nucléotides. Ochoa synthétisa des molécules d'ARN à partir de nucléotides, et Kornberg en fit autant pour l'ADN. Les deux hommes partagèrent le prix Nobel de médecine et de physiologie en 1959. Kornberg montra aussi que si l'on part d'une matrice naturelle d'ADN, l'enzyme catalyse la formation d'une molécule qui est identique à l'ADN naturel. En 1965, Sol Spiegelman, de l'université de l'Illinois, utilisa l'ARN d'un virus (le plus simple des êtres vivants) et produisit de nouvelles molécules similaires. Comme ces nouvelles molécules présentaient les propriétés essentielles du virus, cette expérience constitua l'approche la plus avancée en matière de reproduction de la vie dans un tube à essai. En 1967, Kornberg et d'autres reproduisirent l'expérience, utilisant comme matrice l'ADN d'un virus vivant.

La quantité d'ADN associée à la plus simple manifestation de vie est faible – une seule molécule pour un virus – et peut être réduite. En 1967, Spiegelman procéda à la réplication de l'acide nucléique d'un virus et préleva des échantillons à des intervalles de plus en plus courts, pour tester leur capacité de réplication. Il sélectionna ainsi des molécules qui accomplissaient l'opération à une rapidité inhabituelle, parce qu'elles étaient plus petites que la moyenne. En fin de compte, il avait réduit le virus à un sixième de sa taille normale et multiplié la vitesse de réplication par un facteur quinze.

Bien que ce soit l'ADN qui se réplique dans les cellules, de nombreux virus parmi les plus simples contiennent de l'ARN. Dans ces virus, l'ARN double brin se réplique. Dans les cellules, l'ARN est simple brin et ne se réplique pas.

Cependant, la structure simple brin et la réplication ne s'excluent pas mutuellement. Robert Louis Sinsheimer, un biophysicien américain, découvrit un virus qui contient de l'ADN simple brin. Cette molécule d'ADN doit se répliquer ; mais comment cela peut-il avoir lieu avec un seul brin ? La réponse est aisée. La chaîne unique détermine la production de son propre complément, et le complément permet alors la synthèse du « complément du complément », qui est une copie de la chaîne de départ.

Il est clair que la disposition du simple brin est moins efficace que celle du double brin (ce qui peut éventuellement expliquer pourquoi la première

n'est trouvée que chez certains virus très simples, et la deuxième dans la plupart des autres créatures vivantes). D'abord, parce qu'un simple brin doit se répliquer en deux étapes successives, tandis que les structures à double brin le font en une seule. Ensuite, il semble à présent qu'un seul des brins soit la structure fonctionnelle importante, la lame de la molécule, pour ainsi dire. On peut concevoir le brin complémentaire comme un fourreau pour la lame. Le double brin représente la lame protégée par son fourreau, sauf en action ; le simple brin est la lame toujours exposée et qui est émoussée par accident.

L'ACTIVITÉ DU GÈNE

La réplication ne sert en fait qu'à conserver la molécule d'ADN. Comment cette molécule permet-elle la synthèse d'une enzyme spécifique, c'est-à-dire d'une molécule de protéine spécifique ? Pour la constitution d'une protéine, l'ADN doit diriger la mise en place, dans un ordre donné, d'acides aminés à l'intérieur d'une molécule comprenant des centaines, voire des milliers d'unités. A chaque position, il doit choisir le bon acide aminé parmi vingt autres. Cela serait facile s'il y avait vingt unités correspondantes dans la molécule d'ADN. Mais l'ADN n'est fait que de quatre blocs de construction, les quatre nucléotides. Réfléchissant à ce sujet, l'astronome George Gamow suggéra, en 1954, que les quatre nucléotides soient utilisés en différentes combinaisons, pour constituer ce que nous appelons maintenant *le code génétique* (tout comme le point et le trait du code morse peuvent être combinés de différentes façons pour représenter les lettres de l'alphabet, les chiffres, etc.).

Si l'on prend quatre nucléotides différents (A, G, C, T), deux par deux, il y a 4 × 4, soit 16 combinaisons possibles (AA, AG, AC, AT, GA, GG, GC, GT, CA, CG, CC, CT, TA, TG, TC et TT), ce qui est insuffisant. Si on les prend trois par trois, il y a 4 × 4 × 4, soit 64 combinaisons différentes, plus qu'il n'en faut. On peut s'amuser à dresser une liste des différentes combinaisons et voir si on peut en trouver une soixante-cinquième.

Il semble que chaque *triplet* de nucléotides ou *codon* représente un acide aminé particulier. Au vu du grand nombre possible de codons différents, il pourrait bien y avoir deux ou même trois codons différents représentant un acide aminé donné. Dans ce cas, le code génétique serait ce que les cryptographes appellent *dégénéré.*

Cela soulève deux questions importantes : quel codon (ou codons) correspond à quel acide aminé ? Et comment l'information du codon (qui est soigneusement enfermée dans le noyau où l'on trouve l'ADN) atteint-elle le cytoplasme où a lieu la synthèse des enzymes ?

Répondons d'abord à la deuxième question. Très rapidement, on a pressenti que l'ARN pouvait être le messager, comme les biochimistes français François Jacob et Jacques Lucien Monod le suggérèrent en premier. La structure de cet ARN devrait être très proche de celle de l'ADN, les différences – si elles existent – ne devant pas affecter le code génétique. L'ARN possède du ribose à la place du désoxyribose (un atome d'oxygène supplémentaire par nucléotide) et de l'uracile à la place de la thymine (un groupement méthyl, CH_3, en moins par nucléotide). De plus, l'ARN est présent principalement dans le cytoplasme, et aussi, dans une moindre mesure, dans les chromosomes.

Il n'était pas difficile de voir, puis de démontrer ce qui se passait. A certains moments, quand les deux brins d'ADN enroulés se déroulent, l'un des brins (toujours le même, le tranchant de la lame) se réplique, non pas avec les nucléotides qui forment la molécule d'ADN, mais avec les nucléotides qui constituent les molécules d'ARN. Dans ce cas, une uracile se lie à l'adénine de l'ADN au lieu d'une thymine. La molécule d'ARN, portant l'empreinte du code génétique dans l'arrangement de ses nucléotides, peut alors quitter le noyau et entrer dans le cytoplasme.

Cet ARN a été nommé *ARN messager*, car il transporte le *message* de l'ADN ; sa formule abrégée est *mARN*.

Le biochimiste américain d'origine roumaine George Emil Palade, grâce à des travaux minutieux de microscopie électronique, a démontré en 1956 que des particules minuscules dans le cytoplasme sont le siège de la fabrication des enzymes. Leur diamètre est de 1/2 000 000 de centimètre environ, et elles sont riches en ARN, d'où leur nom de *ribosomes*. On en trouve au moins 15 000 dans une bactérie, et peut-être dix fois plus dans une cellule de mammifère. Ce sont les particules subcellulaires (ou *organelles*) les plus petites. On a déterminé très rapidement que l'ARN messager, transportant le code génétique de par sa structure, se dirige vers les ribosomes où il se fixe sur l'un, puis plusieurs d'entre eux, et que les ribosomes sont le lieu de la synthèse des protéines.

L'étape suivante fut menée à bien par un biochimiste américain, Mahlon Bush Hoagland, qui a coopéré à l'élaboration de la notion de mARN. Il a montré que, dans le cytoplasme, il y a une catégorie de petites molécules d'ARN, *l'ARN soluble ou sARN* : leur petite taille leur permet de se dissoudre librement dans le liquide cytoplasmique.

Approximativement au centre de la molécule, on trouve un triplet de nucléotides particulier qui s'adapte au triplet complémentaire de la chaîne de mARN : si le triplet sur le sARN est AGC, il s'adapte très précisément au triplet UCG de mARN et seulement là. A l'une des extrémités de la molécule de sARN peut se fixer exclusivement un seul et unique acide aminé. Sur chaque molécule de sARN le triplet au centre correspond à un acide aminé particulier qui est fixé à l'une des extrémités de la molécule. Donc, un triplet complémentaire de mARN signifie que seule une certaine molécule de sARN, portant un acide animé donné, se fixera à cet endroit. De nombreuses molécules de sARN pourraient se fixer l'une après l'autre, à la queue leu leu, aux triplets constituant la structure de mARN (les triplets qui ont été moulés sur la molécule d'ADN du gène en question). Tous les acides aminés étant correctement alignés pourraient alors être liés les uns aux autres pour former la molécule d'enzyme.

Comme, par ce moyen, l'information portée par la molécule de mARN est transférée à la molécule protéique de l'enzyme, le sARN est devenu l'*ARN de transfert* ou *tARN*, appellation désormais courante.

Une équipe dirigée par le biochimiste américain Robert William Holley, en 1964, analysa complètement la molécule d'alanine-tARN (l'ARN de transfert qui est lié à l'acide aminé alanine). Cette analyse fut conduite par la méthode de Sanger, en dégradant la molécule en petits fragments à l'aide des enzymes appropriées, ce qui permet d'analyser les fragments et d'en déduire comment ils s'assemblent. L'alanine-tARN, le premier acide nucléique naturel complètement séquencé, est constitué d'une chaîne de soixante-dix-sept nucléotides. En plus des quatre nucléotides habituelle-

ment rencontrés dans les ARN (A, G, C et U), on trouve sept autres nucléotides de structure proche, en plusieurs exemplaires.

On a d'abord supposé que la chaîne simple du tARN se repliait en son milieu, comme une épingle à cheveux, et que les deux extrémités se tressaient en une double hélice. La structure de l'alanine-tARN ne se prête pas à cette théorie. Elle semble plutôt consister en trois boucles, de telle sorte qu'elle ressemble à un trèfle à trois feuilles. Au cours des années suivantes, d'autres molécules de tARN furent analysées en détail, et toutes semblaient posséder la même structure en trèfle à trois feuilles. Holley, grâce à ce travail, partagea le prix Nobel de médecine et de physiologie en 1968.

C'est ainsi que la structure du gène contrôle la synthèse de l'enzyme. Il reste bien sûr de nombreux problèmes à résoudre, car les gènes ne produisent pas continuellement des enzymes à la vitesse maximum. Le gène peut travailler efficacement à certains moments, lentement à d'autres, et quelquefois pas du tout. Certaines cellules fabriquent des protéines à grande vitesse, pouvant combiner au maximum quelque quinze millions d'acides aminés par chromosome et par minute ; d'autres lentement seulement et certaines n'en fabriquent presque pas ; pourtant toutes les cellules d'un même organisme ont la même organisation génétique. En fait chaque cellule de l'organisme est hautement spécialisée, avec des fonctions caractéristiques et une activité chimique propre. Chaque cellule peut synthétiser une protéine donnée rapidement à certains moments et lentement à d'autres. Et pourtant elles ont toutes, à tout moment, la même organisation génétique.

Il est clair que les cellules ont les moyens de bloquer ou de débloquer les molécules d'ADN des chromosomes. A l'aide de ce système de blocage et de déblocage, des cellules différentes ayant la même structure génique peuvent produire différentes combinaisons de protéines et une cellule donnée ayant un arrangement génique constant peut produire différentes combinaisons à différents moments.

En 1961, Jacob et Monod émirent l'hypothèse que chaque gène possédait son propre répresseur, codé par un *gène régulateur*. Ce répresseur, selon sa géométrie qui peut être modifiée par des changements très subtils de l'état des cellules, bloquera ou libérera le gène. En 1967, ce répresseur fut isolé : c'était une petite protéine. Jacob, Monod et André Michael Lwoff reçurent le prix Nobel de médecine et de physiologie en 1965.

A la suite de travaux laborieux effectués depuis 1973, on s'est aperçu que la double hélice longue d'ADN se tord pour former une deuxième hélice (une *superhélice*) autour d'un centre fait d'une chaîne de molécules d'histones, de sorte qu'il se forme une succession d'unités, les *nucléosomes*. Dans ces nucléosomes, certains gènes peuvent être réprimés ou actifs – tout dépend de leur structure fine ; les histones déterminent peut-être quel gène actif est réprimé de temps en temps, ou est activé. Comme d'habitude, les systèmes biologiques paraissent toujours plus complexes que prévu quand on entre dans les détails.

L'information n'est pas entièrement à sens unique, du gène vers l'enzyme. Il existe aussi un contrôle en « retour ». Imaginons un gène codant pour la synthèse d'une enzyme qui catalyse la réaction de conversion de la thréonine (un acide aminé) en un autre acide aminé, l'isoleucine. La présence de l'isoleucine active le répresseur. Celui-ci arrête l'activité du gène qui produit l'enzyme servant à la synthèse de l'isoleucine.

En d'autres termes, quand la concentration d'isoleucine augmente, il y en a moins de formée ; si sa concentration diminue, le gène est débloqué, et il y a une formation accrue d'isoleucine. La machinerie chimique de la cellule (gènes, répresseurs, enzymes et produits terminaux) est extrêmement complexe et entretient des échanges mutuels sophistiqués. On n'est pas prêt de débrouiller complètement ce système.

En attendant, quoi de neuf en ce qui concerne l'autre question : quel codon code pour quel acide aminé ? Les travaux de Marshall Warren Nirenberg et J. Henrich Matthaei, deux biochimistes américains, donnèrent l'amorce d'une réponse en 1961. Ils utilisèrent un acide nucléique synthétique constitué uniquement d'uracile, grâce au système de Ochoa. Cet *acide polyuridylique* était fait d'une longue chaîne de ...UUUUUUUU... et ne pouvait porter que le codon UUU.

Nirenberg et Matthaei ajoutèrent cet acide polyuridylique à un système qui contenait différents acides aminés, des enzymes, des ribosomes, et tous les autres composants nécessaires à la biosynthèse des protéines. De ce mélange ils isolèrent une protéine constituée entièrement de phénylalanine, c'est-à-dire que UUU est équivalent à phénylalanine. Le premier terme du *dictionnaire génétique* était défini.

Ensuite, ils préparèrent un polynucléotide constitué d'une majorité de nucléotides de l'uridine et de petites quantités d'adénine ; on peut alors avoir en plus du codon UUU, un codon UUA, ou AUU ou UAU. Ochoa et Nirenberg montrèrent que dans ce cas-là, la protéine formée contient une majorité de phénylalanines ainsi que quelques rares leucines, isoleucines et tyrosines, trois autres acides aminés.

Lentement, en utilisant des méthodes analogues, le dictionnaire fut complété. On trouva que le code est bien dégénéré, que GAU et GAC peuvent, par exemple, coder pour l'acide aspartique, et que GGU, GGA, GGC et GGG codent tous pour la glycine. On a trouvé en plus des ponctuations. Le codon AUG code pour la méthionine et signifie aussi début de chaîne. C'est, si l'on peut dire, *une lettre majuscule*. De plus, UAA et UAG signalent la fin d'une chaîne ; ce sont des *points*.

En 1967, le dictionnaire était complet (voir tableau 13.1). Nirenberg et son collaborateur, le chimiste américain d'origine indienne Har Gobind Khorana, partagèrent le prix Nobel de médecine et de physiologie (avec Holley) en 1968.

Le décryptage du code génétique n'est pas, cependant, une « fin heureuse » dans le sens où maintenant tout le mystère est éclairci. Il n'y a probablement pas de fin heureuse de ce genre en science, ce qui est une bonne chose, car un univers sans mystère serait horriblement ennuyeux.

On a déterminé le code génétique principalement en faisant des expériences avec les bactéries, dans lesquelles les chromosomes sont bourrés de gènes actifs qui codent pour la formation de protéines. Les bactéries sont des *procaryotes* (du grec signifiant « avant le noyau »), car elles n'ont pas de noyau, mais un matériel chromosomique distribué dans leurs minuscules cellules.

En ce qui concerne les *eucaryotes* qui ont un noyau cellulaire (ce qui inclut toutes les cellules, sauf les bactéries et les algues bleues), la situation est différente. Les gènes *n'y sont pas* entassés de façon compacte sur toute la longueur de la chaîne nucléotidique. Les portions de la molécule d'acide nucléique qui sont utilisées pour la synthèse de mARN et en fin de compte,

Tableau 13.1.

Le code génétique. Dans la colonne de gauche, les symboles des quatre bases de l'ARN (l'uracile, la cytosine, l'adénine et la guanine) représentant la première « lettre » du triplet du codon ; la deuxième lettre est donnée par les symboles placés horizontalement, alors que la troisième lettre, la moins importante, est dans la colonne de droite. Exemple : la tyrosine (Tyr) est codée par UAU ou UAC. Les abréviations des acides aminés sont les suivantes : Phe, phénylalanine – Leu, leucine – Ileu, isoleucine – Met, méthionine – Val, valine – Ser, sérine – Pro, proline – Thr, thréonine – Ala, alanine – Tyr, tyrosine – His, histidine – Glun, glutamine – Aspn, asparagine – Lys, lysine – Asp, acide aspartique – Glu, acide glutamique – Cys, cystéine – Tryp, tryptophane – Arg, arginine – Gly, glycine.

Première lettre			*Deuxième lettre*		*Troisième lettre*
	U	*C*	*A*	*G*	
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	(codon « stop »)	(codon « stop »)	A
	Leu	Ser	(codon « stop » moins fréquent)	Tryp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Glun	Arg	A
	Leu	Pro	Glun	Arg	G
A	Ileu	Thr	Aspn	Ser	U
	Ileu	Thr	Aspn	Ser	C
	Ileu?	Thr	Lys	Arg	A
	Met (« lettre majuscule » – codon d'initiation)	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val (« lettre majuscule »)	Ala	Glu	Gly	G

des protéines (*exons*), sont émaillées de sections (*introns*) que l'on peut qualifier de charabia. Un gène unique qui code pour une enzyme unique sera constitué d'un certain nombre d'exons (partie codante) séparés par des introns (partie non codante), la chaîne nucléotidique se repliant de telle sorte que les exons seront rapprochés pour permettre la synthèse du mARN. L'estimation donnée précédemment dans ce chapitre, de deux millions de gènes dans une cellule humaine, est donc beaucoup trop élevée, si l'on se réfère aux seuls gènes actifs.

La raison pour laquelle les eucaryotes transportent un tel poids mort demeure un mystère. C'est peut-être, tout d'abord, la façon dont les gènes se sont développés ; chez les procaryotes, les introns ont été éliminés pour que les chaînes nucléotidiques soient plus courtes et se répliquent rapidement afin de permettre une croissance et une reproduction plus

rapides. Chez les eucaryotes, les introns ne seraient pas excisés parce qu'ils n'offrent pas d'avantages immédiatement perceptibles. Il n'y a pas de doute que la réponse, quand on l'aura, sera révélatrice.

Pendant ce temps, les scientifiques ont mis au point des méthodes leur permettant de pénétrer au cœur de l'activité des gènes. En 1971, les microbiologistes Daniel Nathans et Hamilton Othanel Smith ont travaillé sur des *enzymes de restriction* qui coupent les chaînes d'ADN de manière spécifique à certaines séquences d'ADN et pas à d'autres. On connaît un autre type d'enzyme, *l'ADN ligase,* qui unit deux chaînes d'ADN. Paul Berg, un biochimiste américain, a coupé des chaînes d'ADN à l'aide d'enzymes de restriction et les a ensuite recombinées différemment par rapport au point de départ. Une molécule d'*ADN recombinant* a ainsi été formée qui ne ressemblait pas à l'original, ni peut-être à aucune ayant jamais existé jusque-là.

Grâce à ces travaux, il devient possible de modifier des gènes ou d'en concevoir de nouveaux ; on peut les faire pénétrer dans une bactérie (ou dans le noyau d'une cellule eucaryote), et donc fabriquer des cellules présentant des propriétés biochimiques nouvelles. Nathans et Smith partagèrent le prix Nobel de physiologie et de médecine en 1978, et Berg celui de chimie en 1980.

Les travaux sur l'ADN recombinant comportent apparemment des risques. Que se passerait-il, si par inadvertance ou délibérément, on fabriquait une bactérie ou un virus capables de produire une toxine contre laquelle l'homme n'aurait aucune défense ? Si un micro-organisme de ce genre s'échappait du laboratoire, il pourrait infliger à l'humanité une épidémie dramatique. En pensant à ces conséquences, Berg et d'autres, en 1974, appelèrent les chercheurs à adhérer à des contrôles stricts des travaux sur l'ADN recombinant.

En fait, à la lumière de l'expérience, on se rendit compte qu'une telle éventualité était fort réduite. On prenait de grandes précautions, les nouveaux gènes introduits dans les micro-organismes produisaient des souches si faibles (il n'est pas si facile de vivre avec un gène *anormal*) qu'elles pouvaient à peine se maintenir en vie dans les conditions les plus favorables.

Par ailleurs, la science de l'ADN recombinant présente de grands avantages. En dehors de l'accroissement des connaissances concernant les détails de la machine cellulaire et des mécanismes de l'hérédité en particulier, elle offre d'autres avantages plus immédiats. Si on modifie un gène de manière appropriée, ou si on insère un gène étranger dans une bactérie, celle-ci peut devenir une usine miniature qui fabrique des molécules dont l'humanité a besoin (plutôt que la bactérie elle-même).

Dans les années quatre-vingts, les bactéries ont été ainsi modifiées pour fabriquer de l'insuline humaine, sous le nom peu attrayant de *humuline.* Bientôt, les diabétiques ne dépendront plus des réserves limitées d'insuline provenant des pancréas de bovins et de porcs, qui est utile mais pas idéale.

D'autres protéines peuvent être produites en modifiant de manière appropriée des micro-organismes : l'interféron et l'hormone de croissance en sont des exemples, et à long terme les possibilités sont illimitées. On ne sera pas surpris que la question se pose maintenant de savoir si de nouvelles formes de vie peuvent être brevetées.

L'origine de la vie

Avec les acides nucléiques, nous touchons aux bases mêmes de la vie. Sans l'ADN, les organismes vivants ne pourraient pas se reproduire, et la vie telle que nous la connaissons n'aurait jamais pu exister. Toutes les substances de la matière vivante – les enzymes et tous leurs produits – dépendent en dernier ressort de l'ADN. Quelle est l'origine de l'ADN, et donc de la vie ?

C'est une question que la science a toujours hésité à poser car l'origine de la vie est liée à des croyances plus fortes encore que ne le sont l'origine de la Terre et de l'Univers. C'est un sujet qui n'est encore abordé qu'avec hésitation et réserve. Le livre du biochimiste russe Alexandre Ivanovitch Oparin, *L'origine de la vie sur la Terre,* a remis le sujet sur le devant de la scène. Dans cet ouvrage, publié en Union Soviétique en 1924, le problème de l'origine de la vie est traité pour la première fois d'un point de vue entièrement matérialiste. Rien de surprenant, puisque l'Union Soviétique n'est pas inhibée par les considérations religieuses des pays occidentaux.

LES PREMIÈRES THÉORIES

La plupart des cultures primitives ont construit des mythes racontant la création des premiers êtres humains (et quelquefois aussi des autres formes de vie) par des dieux ou des démons. Pourtant, elles n'attribuaient pas toujours l'apparition de la vie à une prérogative divine, mais considéraient que les formes moins évoluées de la vie pouvaient surgir spontanément de la matière inerte sans intervention surnaturelle. Par exemple, les insectes et les vers pouvaient provenir de la viande avariée, les grenouilles de la boue, et les souris du blé en voie de putréfaction. Cette idée était fondée sur des observations de la réalité ; la viande avariée, pour prendre l'exemple le plus évident, donne effectivement naissance à des asticots. Il était donc tout à fait naturel d'affirmer que les vers étaient issus de la viande.

Aristote croyait en la *génération spontanée,* ainsi que les grands théologiens du Moyen Age comme saint Thomas d'Aquin et des savants comme William Harvey ou Isaac Newton. L'évidence constatée de visu est, après tout, difficile à réfuter.

Le premier à avoir accepté l'épreuve de l'expérience fut l'Italien Francesco Redi. En 1668, il décida de vérifier si les vers étaient vraiment générés par la viande avariée. Il plaça des morceaux de viande dans une série de bocaux, dont certains étaient recouverts de gaze, et constata que les vers ne se développaient que dans les bocaux découverts, où les mouches avaient libre accès. Redi en conclut que les vers étaient nés d'œufs microscopiques déposés par les mouches et il affirma que, sans les mouches et leurs œufs, la viande ne pourrait jamais produire de vers, quel que soit son état d'avarie ou de putréfaction.

Les résultats de Redi furent confirmés par ses successeurs, et la croyance selon laquelle les êtres vivants provenaient de la matière inerte fut abandonnée. Mais quand on découvrit les microbes, peu après l'époque de Redi, de nombreux scientifiques pensèrent que ces formes de vie, au moins, devaient provenir de la matière inerte. Même dans des bocaux

couverts de gaze, la viande aurait vite fait de se décomposer et de grouiller de micro-organismes. La croyance en la génération spontanée des micro-organismes resta donc très vivace pendant deux siècles encore après les expériences de Redi.

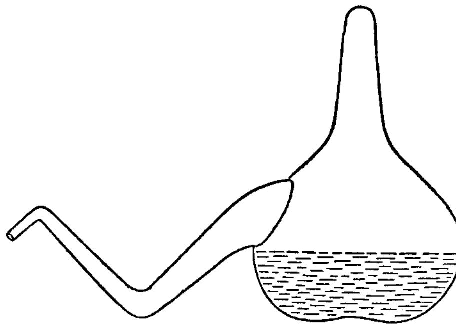
Un autre Italien, Lazzaro Spallanzani, émit le premier des doutes sur cette hypothèse. Il prépara deux séries de flacons contenant du bouillon. Il laissa la première à l'air libre. Il fit bouillir le bouillon de la deuxième série, pour en tuer les organismes déjà présents, et scella les flacons pour prévenir les organismes de l'air ambiant d'y pénétrer. Le bouillon de la première série fut rapidement envahi de micro-organismes ; tandis que le bouillon des récipients stériles restait stérile. Cette expérience apporta la preuve, à la satisfaction de Spallanzani, que même la vie microscopique ne peut naître de la matière inerte. Il isola aussi une bactérie unique et fut le témoin de sa division en deux bactéries.

Les défenseurs de la théorie de la génération spontanée n'étaient pas convaincus. Ils prétendaient que l'ébullition avait détruit quelque *principe vital* et que, donc, aucune vie microscopique ne pouvait se développer dans les flacons fermés contenant le bouillon stérile de Spallanzani. Pasteur résolut cette question une fois pour toutes en 1862. Il conçut un flacon muni d'un long col de cygne en forme de S à l'horizontale (voir figure 13.7). L'ouverture du col permettait à l'air de circuler dans le flacon, mais les particules et les micro-organismes ne le pouvaient pas, car le col courbé servait de piège, comme un siphon d'évier. Pasteur mit du bouillon dans le flacon, le porta à ébullition (pour tuer tous les micro-organismes du col et du bouillon), puis il attendit. Le bouillon restait stérile. Il n'y avait pas de principe vital dans l'air. La démonstration de Pasteur sonna définitivement le glas de la théorie de la génération spontanée.

Le doute persistait néanmoins chez les scientifiques. Comment la vie avait-elle surgi, si ce n'est grâce à la création divine ou à la génération spontanée ?

Vers la fin du XIXe siècle, quelques théoriciens passèrent d'un extrême à l'autre, disant que la vie était éternelle. La théorie qui rencontra le plus grand écho fut celle de Svante Arrhenius (le chimiste qui développa le concept de l'ionisation). En 1907, il publia un livre intitulé *Worlds in the Making* (Le monde en fabrication), dépeignant un univers dans lequel la

Figure 13.7. Le flacon utilisé par Pasteur pour ses expériences sur la génération spontanée.



vie avait toujours existé et migré à travers l'espace, colonisant sans arrêt de nouvelles planètes. La vie voyageait sous forme de spores échappées de l'atmosphère par des mouvements aléatoires, qui étaient conduites à travers l'espace sous la pression de la lumière solaire.

Cette idée, selon laquelle la pression de la lumière est une force conductrice possible, n'a rien de risible. Maxwell a été le premier à prédire l'existence d'une telle pression des radiations sur des bases théoriques et, en 1899, celle-ci fut démontrée expérimentalement par le physicien russe Pierre Nicolaevitch Lebedev.

Arrhenius soutenait que les spores voyageaient continuellement dans l'espace interstellaire, guidées par les radiations lumineuses jusqu'à ce qu'elles meurent ou tombent sur quelque planète ; là, elles se reproduisaient et entraient en compétition avec les formes de vie déjà présentes, ou bien elles communiquaient la vie à la planète inhabitée, mais habitable.

A première vue, cette théorie semble attrayante. Les spores des bactéries, protégées par une capsule épaisse, sont très résistantes au froid et à la déshydratation, et on pourrait concevoir qu'elles survivent longtemps dans le vide de l'espace. Elles ont également la taille adéquate qui les rend plus sensibles à la pression extérieure des radiations solaires qu'à leur propre gravité. Mais la théorie d'Arrhenius est tombée sous les coups de la lumière ultraviolette. En 1910, des chercheurs ont montré que celle-ci tuait rapidement les spores des bactéries ; la lumière ultraviolette du Soleil est très intense dans l'espace interplanétaire, sans parler des autres rayonnements destructeurs, comme les rayons cosmiques, les rayons X solaires et les zones chargées de particules, situées autour de la Terre, telles que la ceinture de Van Allen. On peut imaginer qu'il existe des spores résistantes aux radiations, mais des spores faites de protéines et d'acides nucléiques, telles que nous les connaissons, ne peuvent pas surmonter ces difficultés. Pour s'en assurer, des micro-organismes particulièrement résistants ont été exposés aux radiations de l'espace à bord de la capsule *Gemini 9* en 1966. Elles n'ont survécu que six heures à la violente lumière solaire non filtrée, alors que nous parlons d'expositions non pas de quelques heures, mais de mois et d'années.

D'autre part, si nous supposons que la Terre porte la vie parce qu'elle a été ensemencée par de petits germes provenant d'ailleurs, nous devrions nous demander comment cette vie a surgi ailleurs. Mais nous revenons alors au point de départ.

L'ÉVOLUTION CHIMIQUE

Bien que quelques scientifiques, encore de nos jours, trouvent l'hypothèse de l'ensemencement attrayante, la plus grande majorité pense souhaitable d'expliquer l'origine de la vie ici, sur Terre, par des mécanismes valables.

Les scientifiques repartent de la notion de génération spontanée, mais avec une différence : avant Pasteur, on pensait que la génération spontanée correspondait à quelque chose qui se passait dans le présent et rapidement ; pour l'opinion moderne, cela s'est produit il y a longtemps et fort lentement.

On sait maintenant que la génération spontanée est impossible, car tout ce qui approcherait la complexité nécessaire à la plus élémentaire forme

de vie concevable serait rapidement incorporé, sous forme de nourriture, à l'un des innombrables germes de vie déjà existants. La génération spontanée, donc, ne pouvait avoir lieu que sur une planète où la vie n'existerait pas encore ; dans le cas de la Terre, elle serait apparue il y a plus de trois milliards et demi d'années.

De plus, la vie ne peut pas exister dans une atmosphère riche en oxygène. L'oxygène est un élément actif qui s'associe aux substances au moment où elles s'assemblent pour donner une forme proche de la vie, que l'oxygène dégrade ensuite. Pourtant, comme je l'ai dit au chapitre 5, les scientifiques pensent que l'atmosphère de la Terre, à l'origine, était réductrice et ne contenait pas d'oxygène libre. En fait il est possible que cette atmosphère ait été composée de gaz contenant de l'hydrogène tels que le méthane (CH_4), l'ammoniac (NH_3), la vapeur d'eau (H_20) et peut-être un peu d'hydrogène (H_2).

Nous pourrions appeler atmosphère I une telle atmosphère, à forte teneur en hydrogène. A travers la photodissociation, elle se serait transformée lentement en une atmosphère de gaz carbonique et azote (voir chapitre 5), ou atmosphère II. Ensuite, une couche d'ozone se serait formée dans la partie haute de l'atmosphère et l'évolution spontanée se serait arrêtée. La vie aurait-elle pu alors se former dans l'une ou l'autre de ces atmosphères ?

H.C. Urey pense que la vie a commencé dans l'atmosphère I. En 1952, Stanley Lloyd Miller, alors « thésard » chez Urey, a fait circuler de l'eau, de l'ammoniac, du méthane et de l'hydrogène, et fait passer une décharge électrique dans ce mélange (pour simuler la radiation ultraviolette du Soleil). Au bout d'une semaine, il a analysé la solution par chromatographie sur papier et a trouvé – en plus des substances simples ne contenant pas d'azote – de la glycine et de l'alanine, les deux acides animés les plus simples, ainsi que les précurseurs d'un ou deux autres acides aminés plus complexes.

L'expérience de Miller était significative à plus d'un titre. Tout d'abord, ces substances s'étaient formées rapidement et en quantités étonnamment grandes. Un sixième du méthane avec lequel il avait commencé avait servi à la formation de substances organiques plus complexes ; pourtant, l'expérience n'avait duré qu'une semaine.

Ensuite, le type de molécules organiques formées durant l'expérience de Miller se trouvait précisément être celui existant dans les tissus vivants. La façon dont les molécules simples étaient devenues plus complexes rappelait le processus de la vie. Cette ressemblance a été confirmée par des expériences plus élaborées. A aucun moment on n'a détecté de molécules qui correspondent à un monde incompatible avec la vie.

Philip Abelson a répété les expériences de Miller, en utilisant un matériel constitué de différents gaz dans des proportions différentes. Il s'est avéré que, chaque fois qu'il commençait les expériences avec des molécules contenant des atomes de carbone, d'hydrogène, d'oxygène et d'azote, il obtenait des acides aminés du type que l'on trouve dans les protéines. Les décharges électriques ne furent pas les seules sources d'énergie à donner ces résultats. Ainsi, en 1959, deux chercheurs allemands, Wilhelm Groth et H. von Weyssenhoff, ont réalisé une expérience à base de lumière ultraviolette, et ont obtenu également des acides aminés.

Le fait que, vers la fin des années soixante, on ait trouvé dans les nuages gazeux du cosmos des molécules d'une grande complexité, indique que

les premières étapes du phénomène de la vie ont pu utiliser cette voie (voir chapitre 2). Il se pourrait donc qu'à l'époque où la Terre était formée de nuages de poussières et de gaz, les premières étapes de la formation de molécules complexes aient déjà eu lieu.

La Terre, à sa formation, avait peut-être une provision d'acides aminés. Cette théorie fut mise en évidence en 1970. Le biochimiste ceylanais Cyril Ponnamperuma a étudié une météorite tombée en Australie le 28 septembre 1969. Des analyses minutieuses ont montré les traces de cinq acides aminés – la glycine, l'alanine, l'acide glutamique, la valine et la proline – qui ne présentaient pas d'activité optique. Ils ont donc été formés non pas par des processus vitaux (leur présence n'était pas le résultat d'une contamination terrestre), mais par des réactions chimiques non vitales du genre de celles faites par Miller.

En fait, Fred Hoyle et un collègue indien, Chandra Wickramasinghe, impressionnés par cette découverte, ont pensé que les synthèses pouvaient aller bien au-delà des acides aminés déjà trouvés. Selon eux, des germes microscopiques de vie ont été formés en quantité importante dans l'absolu, mais insuffisante pour qu'ils soient décelés à des distances astronomiques ; ces germes se sont formés non seulement dans des nuages gazeux très éloignés, mais aussi dans des comètes de notre propre système solaire. La vie, de ce fait, serait parvenue sur Terre transportée par les queues des comètes. (Ajoutons que pratiquement personne ne prend cette spéculation au sérieux.)

Les chimistes peuvent-ils dépasser le stade des acides aminés ? L'un des moyens d'y parvenir serait de partir d'échantillons plus grands de matière première et de leur fournir plus d'énergie. Ce procédé créerait des substances de plus en plus nombreuses et de complexité croissante, mais ces mélanges seraient alors trop compliqués pour être analysés facilement.

Les chimistes pourraient aussi partir d'une étape plus avancée, en utilisant les produits obtenus dans les premières expériences comme matière première. Par exemple, l'un des produits de Miller était l'acide cyanhydrique. En 1961, le biochimiste américain d'origine espagnole Juan Oro partit d'un mélange contenant de l'acide cyanhydrique. Il obtint des produits plus riches en acides aminés, et même quelques courts peptides, ainsi que des purines, l'adénine en particulier, l'un des composants essentiels des acides nucléiques. En 1962, Oro employa le formaldéhyde comme matière première et obtint du ribose et du désoxyribose, qui sont aussi des composants des acides nucléiques.

En 1963, Ponnamperuma réalisa des expériences similaires à celles de Miller, utilisant des faisceaux d'électrons comme source d'énergie, et il trouva de l'adénine. Avec Ruth Mariner et Carl Sagan, il poursuivit ces expériences en ajoutant de l'adénine à une solution de ribose ; après exposition à la lumière ultraviolette, il obtint de l'*adénosine*, molécule où l'adénine est liée à du ribose. Si l'expérience avait été faite en présence de phosphate, celui-ci se serait lié aux composés précédents pour former le nucléotide de l'adénine. En effet, la fixation de trois groupements phosphates à l'adénosine produit l'adénosine triphosphate (ATP) (comme c'est expliqué au chapitre 12), qui est une molécule essentielle dans les transferts énergétiques des tissus vivants. En 1965, il forma un *dinucléotide* – deux nucléotides liés. On peut produire d'autres substances si on ajoute des produits – tels que la cyanamide ($CNNH_2$) et l'éthane (CH_3CH_3), qui auraient pu être présents aux origines – aux mélanges employés par les

différents chercheurs dans ce domaine. On peut donc penser que les modifications normales, physiques et chimiques, de l'Océan et de l'atmosphère primordiaux ont agi dans ce sens pour produire des protéines et des acides nucléiques.

On suppose que tout composé formé dans l'Océan sans vie a tendance à s'y maintenir, ne pouvant être consommé ni dégradé par des organismes, grands ou petits. De plus, dans l'atmosphère primordiale, il n'y avait pas d'oxygène libre pour oxyder et couper les molécules. Les seuls facteurs capables de scinder les molécules complexes étaient les ultraviolets et l'énergie provenant de la radioactivité qui ont participé à leur construction. Les courants océaniques ont peut-être transporté la plus grande partie du matériel dans un lieu sûr, à un niveau intermédiaire dans la mer, à mi-distance entre la surface irradiée par les ultraviolets et le fond radioactif. En fait, Ponnamperuma et ses collaborateurs ont estimé qu'au moins 1 % de l'Océan primordial était constitué par ces composés organiques. S'il en était ainsi, ceux-ci représenteraient une masse de plus d'un million de milliards de tonnes ; c'est certainement une quantité suffisante pour que d'une part, les forces naturelles les modifient et que, d'autre part, des substances de la plus haute complexité aient des chances d'être fabriquées dans un court laps de temps (surtout sachant que des milliards d'années étaient disponibles pour cette réalisation).

On peut donc penser que des concentrations plus grandes d'acides aminés complexes et de sucres simples se sont formées avec le temps à partir d'éléments issus de l'Océan et de l'atmosphère primordiaux ; les acides aminés d'une part, les purines, les pyrimidines, les sucres et les phosphates d'autre part, se seraient associés pour former, les uns des peptides, les autres des nucléotides, et ainsi se seraient créés, au fil du temps, des protéines et des acides nucléiques. On en arrive à l'étape clé, la formation, par association au hasard de nucléotides, d'une molécule d'acide nucléique capable de se répliquer. C'est ce moment qui marque le début de la vie.

Ainsi, une période d'*évolution chimique* a précédé l'évolution de la vie elle-même.

Tout laisse penser qu'une seule molécule porteuse de vie a suffi à engendrer la vie et donner naissance au monde vivant, de même qu'une seule cellule fécondée peut donner naissance à un organisme extrêmement complexe. Dans la « purée » organique que constituait l'Océan de ce temps-là, la première molécule porteuse de vie a pu se multiplier à des milliards et des milliards d'exemplaires identiques en peu de temps. Des mutations occasionnelles ont créé des molécules légèrement différentes et celles qui étaient d'une efficacité supérieure se sont multipliées aux dépens de leurs voisines pour remplacer les anciennes. Si un groupe était plus efficace dans l'eau chaude et un autre dans l'eau froide, deux variétés émergeaient, chacune confinée à l'environnement qui lui convenait le mieux. Le cours de l'*évolution organique* aurait été mis en mouvement de cette manière.

Même si, au début, plusieurs molécules porteuses de vie venaient à exister séparément, il est très probable que les plus efficaces supplantaient les autres ; ainsi, toute la vie d'aujourd'hui pourrait très bien descendre d'une seule molécule originelle. En dépit de l'immense diversité des organismes vivants, tous ont le même plan de base fondamental. Leurs métabolismes cellulaires sont presque tous identiques. De plus, il faut

souligner que les protéines de tous les organismes vivants sont composées d'acides aminés de forme L plutôt que de forme D. Il se peut que la nucléoprotéine, d'où toute la vie descend, ait été construite par hasard à partir d'acides aminés de forme L ; comme la forme D ne peut s'associer à la forme L dans aucune chaîne stable, ce qui, au départ, n'était dû qu'au hasard a persisté par multiplication à l'échelle universelle. Ceci ne signifie pas que les acides aminés de forme D sont totalement absents du milieu naturel ; on les retrouve dans les parois cellulaires de quelques bactéries et dans quelques substances antibiotiques, mais ce sont là des cas exceptionnels.

LES PREMIÈRES CELLULES

Bien entendu, il reste une longue distance à parcourir de la molécule jusqu'à la vie telle que nous la connaissons aujourd'hui. A part les virus, toute la vie est organisée en cellules ; et une cellule, aussi petite soit-elle à l'échelle humaine, est extrêmement complexe dans sa structure chimique et ses échanges. Quelle est l'origine des cellules ?

Ce mystère a été éclairci par les recherches du biochimiste américain Sidney Walter Fox. Pour lui la Terre, à ses débuts, devait être très chaude, et la seule énergie de la chaleur aurait suffi à former des composés complexes à partir des plus simples. En 1958, pour vérifier sa théorie, Fox chauffa un mélange d'acides aminés et trouva qu'il formait de longues chaînes qui ressemblaient à celles des molécules de protéines. Ces *protéinoïdes* étaient dégradés par les enzymes qui digèrent les protéines et utilisés comme nourriture par les bactéries.

Fox fit dissoudre les protéinoïdes dans de l'eau chaude et laissa la solution refroidir ; il fut étonné de constater qu'ils s'agglutinaient en formant de petites *microsphères* à peu près de la taille d'une bactérie. Ces microsphères n'étaient pas vivantes selon les critères habituels, mais se comportaient, à certains égards du moins, comme des cellules (elles sont entourées par une sorte de membrane). Fox réussit à faire gonfler et contracter les microsphères, comme de vraies cellules, en modifiant la composition chimique de la solution. Elles produisaient des bourgeons qui semblaient parfois grossir, puis se détacher. Ces microsphères pouvaient aussi se séparer, se diviser en deux ou s'associer pour former des chaînes.

Peut-être existait-il, dans les temps ancestraux, diverses formes de ces agrégats, minuscules et pas tout à fait vivants. Certains étaient particulièrement riches en ADN et doués pour la multiplication, mais pas pour le stockage de l'énergie. D'autres pouvaient transporter l'énergie avec efficacité, mais se répliquaient de manière boiteuse. Enfin des agrégats assemblés en colonies auraient pu coopérer, chacun comblant les carences de l'autre pour former la cellule que nous connaissons, laquelle est beaucoup plus efficace que chacune de ses parties séparées. La cellule moderne a encore un noyau riche en ADN (incapable d'utiliser l'oxygène) et de nombreuses mitochondries qui utilisent l'oxygène, mais ne peuvent se reproduire sans noyau. Possédant encore de petites quantités d'ADN, les mitochondries ont peut-être été des entités indépendantes.

Ces dernières années, on a de plus en plus tendance à penser que l'atmosphère I n'a pas duré très longtemps et que l'atmosphère II a existé pratiquement dès le commencement. Par exemple, les planètes Vénus et

Mars ont des atmosphères II (gaz carbonique et azote). A un moment où elle ne portait pas de vie, comme Mars ou Vénus, la Terre était peut-être aussi entourée d'une atmosphère II.

Ce changement n'est pas catastrophique. Des composés simples peuvent encore être synthétisés à partir du gaz carbonique, de la vapeur d'eau et de l'azote ; l'azote pouvant être convertie en oxydes d'azote, cyanure ou ammoniac par réaction avec le gaz carbonique, l'eau ou les deux, l'énergie étant fournie par les éclairs. Les changements moléculaires progresseraient alors vers la complexité de la vie sous la houlette énergétique du Soleil et des autres sources d'énergie.

LES CELLULES ANIMALES

D'un bout à l'autre de l'existence des atmosphères I et II, les formes primitives de vie n'ont pu exister qu'au prix de la dégradation de substances chimiques complexes en d'autres plus simples, et de la conservation de l'énergie ainsi produite ; les substances complexes étaient reconstruites grâce à l'apport énergétique des ultraviolets solaires. Une fois l'atmosphère II complètement formée et la couche d'ozone mise en place, un risque de pénurie est apparu, la source d'ultraviolets étant coupée. Cependant, à cette époque, quelques agrégats semblables à des mitochondries ont été formés, contenant de la chlorophylle : les ancêtres des chloroplastes. En 1966, les biochimistes canadiens G.W. Hodson et B.L. Baker sont partis du pyrrole et du paraformaldéhyde (deux substances qui peuvent se former à partir de composés simples produits dans des expériences telles que celles de Miller), et ils ont démontré qu'après trois heures de chauffage seulement, des noyaux de porphyrine (la structure de base de la chlorophylle) se formaient.

Au moment où la couche d'ozone s'est formée, l'utilisation même inefficace de la lumière visible par les premiers agrégats contenant de la chlorophylle était malgré tout préférable à la lente asphyxie des systèmes qui en étaient dépourvus. La lumière visible pouvait traverser aisément la couche d'ozone et son énergie plus faible, comparée à la lumière ultraviolette, était suffisante pour activer les systèmes à base de chlorophylle.

Les premiers organismes utilisant la chlorophylle étaient peut-être aussi simples que les chloroplastes actuels. Par exemple, les *algues bleues* (appelées ainsi d'après la couleur des premiers exemplaires étudiés, bien que toutes ne soient pas bleues), dont on connaît 2 000 espèces, sont des organismes unicellulaires vivant de la photosynthèse : des cellules procaryotes très simples, avec une structure semblable à celle des bactéries, sauf qu'elles contiennent de la chlorophylle. Les algues bleues sont peut-être les plus simples descendants des chloroplastes ancestraux, alors que les bactéries seraient des descendants des chloroplastes qui ont perdu leur chlorophylle, sont devenus des parasites ou vivent de la décomposition des tissus morts.

Les chloroplastes se multipliant dans les mers anciennes, le gaz carbonique a été progressivement consommé et remplacé par l'oxygène moléculaire formant l'actuelle atmosphère III. Les cellules végétales sont devenues plus efficaces, chacune contenant de nombreux chloroplastes. A la même époque, faute de nourriture nouvellement créée dans l'Océan, sauf dans les cellules végétales, les cellules complexes ne contenant pas

de chlorophylle ne pouvaient pas exister. Malgré tout, des cellules sans chlorophylle mais contenant des mitochondries à l'équipement sophistiqué pouvaient traiter, avec une grande efficacité, des molécules complexes et stocker l'énergie libérée lors de leur dégradation. Ces cellules pouvaient vivre de l'ingestion de cellules végétales les dépouillant des molécules que ces dernières avaient laborieusement construites. Ainsi sont nées les cellules animales. Finalement, les organismes sont devenus assez complexes pour laisser les fossiles (végétaux et animaux) que nous connaissons aujourd'hui.

Pendant ce temps, l'environnement de la Terre changeait fondamentalement, du point de vue de la création de vies nouvelles. L'évolution chimique ne pouvait plus permettre la naissance ni le développement de formes de vie nouvelles. Tout d'abord, les sources d'énergie qui les avaient engendrées (les ultraviolets et l'énergie de la radioactivité) n'existaient plus ou du moins avaient sérieusement diminué. Ensuite, les formes de vie déjà établies consommaient rapidement toute nouvelle molécule organique, spontanément apparue. Ces deux raisons font qu'il n'y a aucune chance pour qu'une forme de vie indépendante jaillisse du néant (en mettant de côté la possibilité d'une intervention humaine, au cas où nous trouverions l'astuce). De nos jours, la génération spontanée est si improbable qu'elle peut être considérée comme impossible.

La vie dans les autres mondes

Si nous acceptons l'idée que la vie a jailli simplement par le jeu des lois physiques et chimiques, il s'ensuit en toute vraisemblance que la vie n'est pas limitée à la Terre. Quelles sont les possibilités de vie ailleurs, dans l'Univers ?

Lorsque l'on reconnut que les planètes du système solaire étaient des mondes, il fut admis qu'ils étaient les sièges de la vie, même de la vie intelligente. Ce fut avec un certain étonnement que l'on découvrit qu'il n'y avait ni eau, ni air sur la Lune, et donc pas de vie.

Dans le monde actuel de fusées et de sondes (voir chapitre 3), les scientifiques sont pratiquement convaincus qu'il n'y a pas de vie à l'intérieur du système solaire, excepté sur la Terre.

Il en va de même, semble-t-il, à l'extérieur du système solaire. Jupiter a une atmosphère épaisse et complexe, avec une température très basse au niveau de la couche de nuages visible, et très élevée à l'intérieur. A une profondeur et à une température modérées, et du fait de la présence connue de l'eau et de composés organiques, on peut imaginer (comme le suggère Carl Sagan) que la vie y existe. Si c'est vrai pour Jupiter, ce pourrait être vrai aussi pour les trois autres géants gazeux.

Europe, le satellite de Jupiter, est complètement ceinturé par un glacier, sous lequel il pourrait y avoir un océan chauffé sous l'influence de l'attraction de Jupiter. Titan, comme Triton, le satellite de Neptune, a une atmosphère de méthane et d'azote et posséderait aussi de l'azote liquide et des composés organiques solides à la surface. Sur ces trois satellites, il est concevable qu'une certaine forme de vie puisse exister.

Cependant, ce sont des spéculations à long terme. Dans l'état actuel des choses, nous ne pouvons guère espérer en savoir plus ; contentons-nous de supposer – pour ce qui est du système solaire – que la vie ne peut prospérer que sur Terre, nulle part ailleurs. Mais il n'y a pas que le système solaire. La vie peut-elle exister ailleurs dans l'Univers ?

Le nombre total d'étoiles connues dans l'Univers est estimé à environ dix milliards de billions. Notre seule galaxie contient au moins deux cents milliards d'étoiles. Si toutes les étoiles proviennent du même processus que celui que nous croyons être à l'origine de notre système solaire (c'est-à-dire la condensation d'un grand nuage de poussière et de gaz), alors il est vraisemblable qu'aucune étoile n'est isolée, mais que chacune fait partie d'un système local contenant plus d'un corps stellaire. Nous savons qu'il y a beaucoup d'étoiles doubles tournant autour d'un centre commun, et l'on suppose qu'au moins une moitié d'entre elles appartient à un système d'au moins deux étoiles.

Dans l'idéal, ce que nous cherchons, c'est un système multiple dans lequel une partie des membres sont trop petits pour être lumineux par eux-mêmes, et qui sont des planètes plutôt que des étoiles. Bien que nous n'ayons à présent aucun moyen de détecter les planètes au-delà de notre système solaire, même les étoiles les plus proches, nous pouvons en recueillir la preuve indirectement. Cette expérience a été réalisée à l'observatoire Sproul du collège Swarthmore, sous la direction de l'astronome américain d'origine néerlandaise, Peter Van de Kamp.

En 1943, certaines irrégularités de l'une des étoiles de la double étoile 61 de la constellation du Cygne indiquaient la présence d'un troisième composant, trop petit pour être lumineux par lui-même. Ce troisième composant, 61 Cygni C, aurait environ huit fois la masse de Jupiter et, de ce fait (en supposant une même densité), environ deux fois son diamètre. En 1960, une planète d'une taille similaire était localisée tournant autour de la petite étoile Lalande 21185 (localisée par déduction pour expliquer les irrégularités du mouvement de cette étoile). En 1963, une étude détaillée de l'étoile de Barnard indiquait aussi la présence d'une planète à cet endroit, n'ayant qu'une fois et demi la taille de Jupiter.

L'étoile de Barnard est la deuxième constellation la plus proche de nous ; Lalande 21185, la troisième et 61 Cygni, la douzième. Que trois systèmes planétaires existent si près de nous est tout à fait improbable, sauf si les systèmes planétaires sont très répandus. Les étoiles se trouvent à des distances si lointaines que seules les plus grosses planètes peuvent être détectées et ce, avec difficulté. Là où existent des planètes plus grandes que Jupiter, on peut penser qu'il en existe aussi de plus petites.

Malheureusement, les observations qui ont conduit à supposer que les planètes extrasolaires existent ne sont pas très précises, à la limite de l'observable. Les astronomes doutent que l'existence de ces planètes ait été vraiment démontrée.

Cependant, récemment, de nouvelles preuves ont été apportées à cette hypothèse. En 1983, un satellite astronomique à infrarouges (IRAS) a été mis en orbite autour de la Terre. Il était conçu pour détecter et étudier les sources infrarouges dans le ciel. En août, les astronomes Harmut H. Aumann et Fred Gillett ont orienté le système de détection du satellite en direction de l'étoile Véga et trouvé, à leur grande surprise, que Véga est bien plus brillante dans la lumière infrarouge qu'on ne s'y attendait.

Une étude plus détaillée a montré que la radiation infrarouge ne provenait pas de Véga elle-même, mais de son environnement immédiat.

Apparemment, Véga est entourée d'un nuage de matière dont le diamètre est deux fois supérieur à l'orbite de Pluton autour du Soleil. Le nuage est probablement constitué de particules plus grandes que des grains de poussière (sinon ils auraient été aspirés depuis longtemps par Véga). Véga est beaucoup plus jeune que notre Soleil, car elle a moins d'un milliard d'années, elle est soixante fois plus lumineuse que le Soleil et a un vent stellaire plus fort qui empêche les particules de se regrouper. Pour toutes ces raisons, on peut penser qu'un système planétaire est en cours de formation autour de Véga. Dans son vaste nuage de graviers pourraient se trouver des objets déjà planétisés.

Cette découverte vient confirmer l'hypothèse que les systèmes planétaires sont communs dans l'Univers, peut-être aussi communs que les étoiles.

Mais, même en supposant que toutes ou la plupart des étoiles ont des systèmes planétaires, et que beaucoup de planètes sont de la taille de la Terre, nous devons connaître les critères que ces planètes doivent remplir pour être habitables. Le scientifique américain, spécialiste de l'espace, Stephen H. Dole, a étudié ce problème dans son livre *Habitable Planets for Man* (planètes habitables par l'homme, 1964), et a tiré certaines conclusions qu'il reconnaît être des spéculations, mais qui sont plausibles.

Il a mis en évidence, en premier lieu, qu'une étoile doit être d'une certaine taille pour posséder une planète habitable. Plus l'étoile est grande, plus courte est sa durée de vie ; si elle dépasse une certaine taille, elle ne vivra pas assez longtemps pour permettre à une planète de traverser les longues étapes de l'évolution chimique préalable au développement de formes de vie complexes. Une étoile trop petite ne peut pas chauffer suffisamment une planète, à moins que celle-ci soit si proche qu'elle souffrira des effets préjudiciables de l'attraction de l'étoile. Dole en a conclu que seules les étoiles des classes spectrales F2 à K1 sont aptes à posséder des planètes où des êtres humains puissent vivre et que nous puissions coloniser sans trop d'efforts (si les voyages interstellaires devenaient accessibles). Il existe, selon Dole, dix-sept milliards d'étoiles de cette sorte dans notre galaxie.

L'étoile idéale pourrait ainsi satisfaire à tous les critères nécessaires à la présence d'une planète habitable mais ne pas en posséder obligatoirement. Dole a estimé les probabilités qu'une étoile de taille appropriée ait une planète de masse voulue, située à la bonne distance, avec une période de rotation et une orbite correctes. Sa conclusion fut que six cents millions de planètes habitables existeraient dans notre galaxie, chacune contenant déjà quelque forme de vie.

En supposant que ces planètes habitables soient réparties plus ou moins régulièrement dans la Galaxie, Dole a estimé qu'il y a une planète habitable par 80 000 années-lumière au cube. La planète habitable la plus proche serait à quelque vingt-sept années-lumière de distance, et à une centaine d'années-lumière d'ici, il pourrait y avoir une cinquantaine de planètes habitables.

Dole a dressé la liste de quatorze étoiles, situées à moins de vingt-deux années-lumière de chez nous, qui pourraient posséder des planètes habitables. Pour chacune des étoiles, il a évalué la probabilité de cette présence. Il en a déduit que les planètes habitables devaient se trouver

plus probablement parmi les étoiles les plus proches de nous : les deux étoiles ressemblant à des soleils du système Alpha Centauri (Alpha Centauri A et B). Ces deux étoiles regroupées ont, d'après Dole, une chance sur dix de posséder des planètes habitables. La probabilité totale des quatorze étoiles les plus proches de posséder une planète habitable serait donc d'environ 2 sur 5.

Si la vie est la conséquence des réactions chimiques décrites dans les pages précédentes, son développement serait inévitable, sur toute planète semblable à la Terre. Bien entendu, on peut trouver la vie sur une planète, mais pas une vie intelligente. Nous n'avons pas de moyens de faire de suppositions quant à la vraisemblance du développement de l'intelligence sur une planète ; Dole, par exemple, a eu la prudence de ne pas en faire. Après tout, notre Terre, la seule planète habitable que nous puissions étudier, a existé plus de trois milliards d'années sous une forme de vie qui n'était pas une vie intelligente.

Il est possible que les marsouins et certains de leurs semblables soient intelligents, mais n'ayant pas de membres, ils n'ont pas pu inventer l'usage du feu ; par conséquent, leur intelligence, si elle existe, n'a pas pu s'appliquer au développement de technologies. Si l'on considère seulement la vie sur terre, alors ce n'est que depuis un million d'années, ou presque, que l'on peut s'enorgueillir d'y trouver une créature d'une intelligence supérieure à celle du singe.

Cela signifie qu'il n'y a eu de vie intelligente que pendant 1/3 500 du temps où il y a eu une forme de vie sur la Terre (approximation grossière). Si nous pouvons dire que de toutes les planètes porteuses de vie, une sur 3 500 porte une vie intelligente, alors sur les 600 millions de planètes habitables estimées par Dole, il y en aurait 180 000 abritant une vie intelligente. Nous pourrions bien ne pas être les seuls dans l'Univers.

Cette conception d'un Univers riche en formes d'intelligence, que Dole, Sagan (et moi-même) partageons, ne fait pas l'unanimité des astronomes. Vénus et Mars ayant été étudiées dans le détail et trouvées hostiles à la vie, une opinion pessimiste s'est fait jour, selon laquelle la vie ne se formerait et ne se maintiendrait pendant des milliards d'années que dans des conditions très restrictives, et que la Terre a beaucoup de chances de remplir ces conditions. Une faible modification de ces conditions, et la vie ne se serait pas formée ; si elle s'était formée, elle n'aurait pas duré.

Dans cette optique, il se pourrait qu'il n'y ait que deux ou trois planètes vivables par galaxie, et seulement une seule ou deux civilisations techniquement avancées dans tout l'Univers.

Selon Francis Crick, il y aurait un nombre mesurable de planètes dans chaque galaxie qui seraient habitables, mais qui ne rempliraient pas les conditions nécessaires à la naissance de la vie. Il se pourrait alors que la vie naisse sur une planète donnée et que, une fois la civilisation arrivée au stade des vols interstellaires, la vie s'établisse ailleurs. La Terre n'a pas encore réalisé de vols interstellaires, mais peut-être des voyageurs sont-ils venus il y a des milliards d'années, qui ont transmis la vie à la Terre, volontairement ou pas.

Ces deux points de vue – l'optimiste et le pessimiste – d'un Univers plein de vie ou presque vide – sont des *conjectures.* Les deux supposent un raisonnement à partir de certaines hypothèses, aucune n'étant étayée par des observations.

Peut-on obtenir des preuves ? Y a-t-il moyen d'affirmer que la vie existe quelque part au voisinage d'une étoile éloignée ? On peut penser que toute

forme de vie suffisamment intelligente pour être arrivée au stade d'une civilisation de haute technologie, comparable ou supérieure à la nôtre, aurait certainement inventé la radioastronomie et pourrait alors transmettre des signaux radio ou les envoyer par inadvertance.

Les scientifiques américains ont pris au sérieux cette hypothèse et, sous la conduite de Frank Donald Drake, ils ont conçu le projet Ozma (nom dérivé du livre d'*Oz* pour les enfants), destiné à enregistrer d'éventuels signaux radio provenant d'autres planètes. Il s'agit de rechercher un certain type d'ondes radio émanant de l'espace. S'ils détectaient des signaux dans un système complexe mais organisé – à l'opposé de ceux qui sont aléatoires, provenant d'étoiles ou de matière excitée dans l'espace, ou de la simple périodicité des pulsars –, on pourrait alors envisager de tels signaux comme des messages d'une intelligence extra-terrestre. Bien sûr, même si ces messages parvenaient jusqu'à nous, la communication avec cette intelligence éloignée resterait un problème. Les messages seraient émis depuis bien des années et les réponses mettraient aussi beaucoup de temps à atteindre les destinataires, puisque la planète la plus proche potentiellement habitable est située à 4,33 années-lumière.

Les sections d'écoute du ciel ont été orientées durant le projet Ozma dans les directions d'Epsilon Eridani, Tau Ceti, Omicron 2 Eridani, Epsilon Indi, Alpha Centauri, 70 Ophiuchi et 61 Cygni. Après deux mois de résultats négatifs, le projet a été arrêté.

D'autres tentatives similaires ont été encore plus brèves et moins élaborées. Les scientifiques ne désespèrent pas malgré tout de faire mieux.

En 1971, un groupe de la NASA conduit par Bernard Oliver a proposé ce qui allait devenir le projet Cyclope. Celui-ci serait constitué d'une grande batterie de radiotélescopes de cent mètres de diamètre chacun, tous disposés en formation, tous dirigés simultanément par un système électronique informatisé. L'ensemble équivaudrait à un seul radiotélescope de quelque dix kilomètres de diamètre. Cet ensemble de radiotélescopes serait capable de détecter des faisceaux radio, comme ceux que la Terre laisse échapper par inadvertance, provenant de distances de cent années-lumière, et pourrait capter une onde de radio balise délibérément dirigée par une autre civilisation à une distance de mille années-lumière.

Pour mettre sur pied un tel ensemble, il faudrait vingt ans et cela coûterait cent milliards de dollars. Avant de s'exclamer devant cette somme, il faut savoir que le monde dépense cinq cents milliards de dollars, c'est-à-dire cinq fois plus, *chaque année,* pour ses dépenses en armements.

Mais pourquoi faire l'essai ? Il y a peu de chances que nous réussissions et dans ce cas, à quoi bon ? Y a-t-il quelque espoir que nous comprenions le message interstellaire ? Pourtant, il y a des raisons d'essayer.

Tout d'abord, le fait d'essayer fera progresser l'art de la radiotélescopie et la compréhension de l'Univers, pour le plus grand avantage de l'humanité. Ensuite, si nous cherchons des messages venant du ciel et n'en recevons aucun, nous pouvons cependant y trouver un intérêt. Mais, si nous détections effectivement un message et que nous ne le comprenions pas ? Qu'est-ce que cela nous ferait ?

Un argument supplémentaire plaide contre l'existence de la vie intelligente sur d'autres planètes : si elle existe et qu'elle nous soit supérieure, pourquoi n'avons-nous pas été découverts ? La vie existe sur Terre depuis des milliards d'années sans qu'elle ait été perturbée par des

influences extérieures (autant qu'on puisse le dire), et c'est ce qui tendrait à prouver qu'il n'existe pas d'influence extérieure.

D'autres arguments peuvent être utilisés pour contrer celui-ci. Peut-être les civilisations qui existent ailleurs sont-elles à une telle distance qu'on ne peut pas nous atteindre, qu'aucun voyage interstellaire n'a jamais été réalisé par une civilisation, que nous sommes tous tellement isolés que nous ne pouvons nous joindre que par des messages longue distance. Peut-être aussi qu'ils nous *ont* atteints, mais qu'ils se sont rendu compte que nous sommes une planète en voie de développement vers la civilisation, et que donc ils s'abstiennent d'intervenir.

Ce sont des arguments de faible portée. Il y en a un plus puissant et plus effrayant. Il se pourrait que l'intelligence soit une propriété autodestructrice. Peut-être que dès qu'une espèce atteint un niveau de technicité élevé, elle se détruit, comme nous semblons le faire nous aussi avec nos stocks grandissants d'armes nucléaires, notre penchant au surpeuplement et à la destruction de l'environnement. Dans ce cas, l'absence de civilisation n'est pas en cause : il y en a peut-être de nombreuses qui ne sont pas encore arrivées au stade d'envoyer et de recevoir des messages. Il n'en reste peut-être qu'une ou deux qui ont survécu à cette autodestruction et qui sont sur le point d'envoyer des messages avant de s'autodétruire.

Au cas où nous recevrions *un* message nous révélant qu'il existe quelque part *une* civilisation ayant atteint un haut niveau de technicité (un niveau plus élevé que le nôtre, en toute vraisemblance) et qui ne se serait *pas* détruite, c'est qu'elle se serait débrouillée pour survivre, et donc pourquoi pas nous ?

C'est le genre d'encouragement dont l'humanité a grand besoin à ce stade de son histoire et un message que, moi du moins, je serais heureux de recevoir.

Chapitre 14

Les micro-organismes

Les bactéries

Jusqu'au XIIe siècle, on ne connaissait pas d'êtres vivants de taille inférieure aux tout petits insectes et on ne pensait pas qu'il existât d'organismes plus petits. On aurait probablement admis que certains êtres vivants puissent être rendus invisibles par des pouvoirs surnaturels, mais non qu'il y ait dans la nature des êtres vivants invisibles par leur taille.

LES INSTRUMENTS DE GROSSISSEMENT OPTIQUE

Si on avait soupçonné l'existence de ces êtres vivants, les instruments de grossissement optique seraient sans doute apparus plus tôt dans l'histoire. En effet, les Grecs et les Romains savaient déjà qu'il était possible de concentrer la lumière solaire en un point au moyen d'objets de verre d'une certaine forme (comme par exemple une sphère de verre creuse remplie d'eau) et de grossir ainsi les objets vus à travers le verre. Ptolémée, par exemple, s'était intéressé à l'optique des loupes et ses observations avaient été exposées plus en détail par des savants arabes tels que Alhazen, en l'an 1000 apr. J.-C.

C'est au début du XIIIe siècle qu'un évêque anglais, Robert Grosseteste, également philosophe et scientifique amateur passionné, proposa d'utiliser les verres optiques appelés *lentilles* (dont la forme ressemble à celle des graines du même nom) dans le but d'agrandir les objets trop petits pour être facilement observés à l'œil nu. L'un de ses élèves, Roger Bacon, appliqua cette idée à la fabrication de lunettes de correction pour la vue.

Au début, on ne fabriqua que des verres convexes qui corrigeaient la vision de loin, et ce n'est que vers 1400 que les verres concaves firent leur apparition. L'invention de l'imprimerie fit beaucoup augmenter les demandes en lunettes ; et vers le XVIe siècle, la lunetterie était devenue une profession artisanale, particulièrement aux Pays-Bas, dont c'était la grande spécialité.

Les lunettes bifocales, utilisées pour améliorer à la fois la vision de près et de loin, furent inventées par Benjamin Franklin en 1760 ; les premiers verres destinés à corriger l'astigmatie furent conçus par l'astronome anglais George Biddell Airy, lui-même astigmate, en 1827 ; et on doit l'idée des lentilles de contact à un médecin allemand, Adolf Eugen Fick, en 1887.

Retournons maintenant aux fabricants de lunettes hollandais. On raconte qu'en 1608, un apprenti travaillant chez un lunetier nommé Hans Lippershey fit un jour une constatation étonnante : alors qu'il occupait ses heures de liberté à regarder des objets au travers de deux lentilles superposées, il remarqua que quand il les tenait à une certaine distance l'une de l'autre, les objets éloignés paraissaient plus proches. L'apprenti rapporta aussitôt cette observation à Lippershey, qui se mit à construire le premier télescope en plaçant les deux lentilles dans un tube pour les maintenir correctement espacées. La valeur militaire de cet instrument fut vite comprise par le commandant des forces hollandaises dans la rébellion contre l'Espagne, le prince Maurice de Nassau, qui conseilla de garder la découverte secrète.

C'était compter sans Galilée, qui avait eu vent de l'invention d'un instrument optique rapprochant les objets éloignés et qui, sans rien savoir de plus que le fait qu'il était construit avec des lentilles, en découvrit bientôt le principe et construisit son propre télescope, lequel fut terminé six mois après celui de Lippershey.

Galilée s'aperçut aussi qu'en réarrangeant les lentilles de son télescope, il pouvait agrandir des objets proches, ce qui en faisait en fin de compte un *microscope.* Pendant les décennies qui suivirent, plusieurs scientifiques construisirent des microscopes, en particulier un naturaliste italien du nom de Francesco Stelluti qui étudia l'anatomie des insectes, Malpighi qui fit la découverte des capillaires du rein, et Hooke à qui on doit des observations sur les cellules du liège.

Mais ce ne fut que lorsqu'Anton Van Leeuwenhoek, un marchand de la ville de Delft, s'intéressa au microscope que celui-ci prit sa valeur réelle ; en effet, certaines des lentilles de Van Leeuwenhoek pouvaient agrandir une image par un facteur 200.

Van Leeuwenhoek se servit de son microscope pour observer des objets de toutes sortes, décrivant en détail ce qu'il voyait dans des lettres à la Royal Society de Londres. Et lorsqu'il fut élu membre de cette organisation si aristocratique, lui qui n'était qu'un artisan, ce fut un véritable triomphe pour la démocratie des milieux scientifiques. L'humble fabricant de microscopes de Delft reçut même de son vivant la visite de la reine d'Angleterre et du tsar Pierre le Grand.

A travers ses verres, Van Leeuwenhoek put décrire l'aspect des cellules du sperme et observer la circulation du sang dans les capillaires de la queue d'un têtard. Qui plus est, il fut le premier à apercevoir, en 1675, des *animalcules* vivant dans l'eau stagnante, des êtres vivants trop petits pour être vus à l'œil nu. Il observa aussi les toutes petites cellules de la levure, et découvrit en 1676, à la limite du pouvoir grossissant de ses lentilles, les *germes* * qu'aujourd'hui nous appelons *bactéries.*

Ensuite, les microscopes ne s'améliorèrent que lentement et cela prit un siècle et demi avant que des objets de la taille des microbes puissent être étudiés sans peine. Le *microscope achromatique,* par exemple, (dans lequel la distance focale est la même pour tous les rayons lumineux quelle que soit leur couleur, ce qui améliore la netteté de l'image) ne fut inventé qu'en 1830 par l'opticien anglais Joseph Jackson Lister qui s'en servit pour observer les globules rouges, d'abord détectés comme des taches sans contours nets par le médecin hollandais Jan Swammerdam en 1658, et qui étaient en fait des disques biconcaves – comme de petits anneaux avec un renfoncement au lieu d'un trou. Le microscope achromatique constitua un grand progrès ; et en 1878, le physicien allemand Ernst Abbe y ajouta une série d'améliorations qui aboutirent à ce que l'on peut appeler le microscope optique moderne.

LA NOMENCLATURE BACTÉRIENNE

Ce monde microscopique tout neuf contenait de nombreuses formes de vie différentes qui furent identifiées petit à petit. On sait maintenant que les animalcules de Van Leeuwenhoek font partie du règne animal, se nourrissent de particules plus petites et se déplacent par divers moyens (*flagelles, cils,* ou avancées de protoplasme appelées *pseudopodes*). Ces animaux, dont le caractère unicellulaire a été établi par le zoologue allemand Karl Ernst Siebold, s'appellent des *protozoaires* (du grec pour « premiers animaux »).

Les microbes sont des êtres beaucoup plus petits et plus simples que les protozoaires. En dehors de quelques exceptions, ils ne bougent pas, et leur principale caractéristique est de croître et de se multiplier ; comme ils ne montrent pas les propriétés typiques du règne animal, on les a souvent classés parmi les *moisissures* – des végétaux sans chlorophylle qui se nourrissent de matières organiques. Presque tous les biologistes à l'heure actuelle les considèrent comme faisant partie d'un règne autonome, et non pas du règne animal ou végétal.

C'est en 1773 que le microscopiste danois Otto Frederik Müller, en observant ces êtres microscopiques en détail, put distinguer deux catégories : les *bacilles* (du mot latin signifiant « bâtonnets ») et les *spirilles* (qui ont l'aspect d'une hélice). Grâce au microscope achromatique, le chirurgien autrichien Theodor Billroth décrivit des variétés encore plus petites qu'il désigna du nom de *coccus* (du mot grec pour « baie »). Et ce

* Ce terme de *germes* ne me semble pas très approprié car il est employé pour d'autres usages, en particulier dans « germe de blé » pour la partie vivante d'une graine, dans « cellules germinales » et « lignée de cellules germinales » pour les cellules sexuelles et les organes embryonnaires, et de façon générale pour ce qui possède le potentiel de la vie. *(N.D.T.)*

fut finalement le botaniste allemand Ferdinand Julius Cohn qui inventa le nom de *bactérie* (dérivé d'un autre mot latin signifiant « bâtonnet ») (voir figure 14.1).

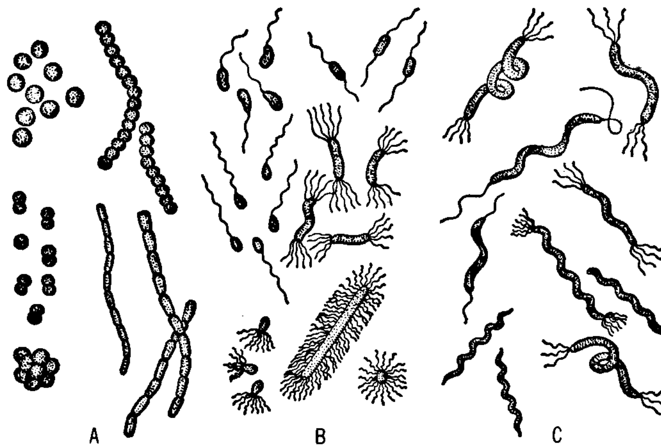
Pasteur utilisait le mot *microbe* (« petite vie ») pour désigner toutes les formes de vie microscopiques – végétales, animales et bactériennes ; mais comme ce mot changea rapidement de sens pour n'être attribué qu'aux bactéries récemment découvertes, c'est le terme de *micro-organismes* que l'on emploie aujourd'hui pour désigner toutes ces formes de vie.

On peut considérer que les cellules des organismes eucaryotes multicellulaires du règne animal et végétal (y compris les cellules du corps humain) sont les plus grands des micro-organismes. Cependant les cellules des protozoaires, qui possèdent elles aussi des noyaux, des mitochondries et autres organelles, sont souvent plus grandes et plus complexes que celles, par exemple, du corps humain. Ceci est dû au fait que la cellule protozoaire doit accomplir toutes les fonctions inséparables de la vie, tandis que les cellules des organismes multicellulaires peuvent se spécialiser et dépendre d'autres cellules pour accomplir certaines fonctions ou leur fournir les produits qu'elles-mêmes ne peuvent apporter.

Les cellules végétales unicellulaires, appelées *algues,* sont pour les mêmes raisons aussi complexes que les cellules des plantes multicellulaires et contiennent des noyaux, des chloroplastes, etc.

Par contre, les bactéries sont des organismes procaryotes qui ne contiennent ni noyau, ni autres organelles. Leur matériel génétique est dispersé à travers la cellule, alors que chez les eucaryotes il est contenu à l'intérieur d'un noyau. Les bactéries se caractérisent aussi par la présence d'une paroi cellulaire constituée de substances complexes combinant des polysaccharides et des protéines. Enfin, leur taille (leur diamètre varie entre un et dix microns) est de beaucoup inférieure à celle des cellules eucaryotes.

Figure 14.1. Les différents types de bactéries : les cocci (A), les bacilles (B) et les spirilles (C). Chaque type comporte un certain nombre de variétés.



Les *algues bleues* constituent un autre grand groupe de procaryotes dont les principales différences avec les bactéries sont la présence de chlorophylle et l'aptitude à la photosynthèse ; le terme d'algues étant d'ailleurs plutôt réservé aux végétaux unicellulaires eucaryotes.

Les bactéries ne sont simples qu'en apparence. Bien qu'elles n'aient pas de noyau, et ne semblent pas transférer leur chromosome selon un mode de reproduction sexuelle classique, elles ont néanmoins une forme primitive de sexualité. C'est en 1946 qu'Edward Tatum et son étudiant Joshua Lederberg constatèrent que les bactéries transfèrent parfois des morceaux d'acides nucléiques d'une cellule à l'autre. Lederberg appela ce procédé *conjugaison* et, à la suite de cette découverte, lui et Tatum partagèrent en 1958 le prix Nobel de physiologie et de médecine.

En étudiant la conjugaison, on s'aperçut que les morceaux d'acides nucléiques qui subissaient le transfert étaient des molécules circulaires – et non linéaires. Ces cercles d'acides nucléiques, appelés *plasmides* par Lederberg en 1952, sont ce que les bactéries ont de plus semblable aux organelles des eucaryotes ; ces plasmides ont en effet un génôme, contrôlent la synthèse de certaines enzymes, et peuvent transférer des propriétés biologiques d'une cellule à une autre.

LES MALADIES MICROBIENNES

C'est à Pasteur que l'on doit la découverte d'une relation entre micro-organismes et maladies, découverte qui est à la base de la science moderne de la *bactériologie* (appelée aussi *microbiologie*). C'est un problème industriel, plutôt que médical, qui l'amena à formuler cette théorie. Dans les années 1860-1870, l'industrie française de la soie était gravement affectée par une maladie du ver à soie, et elle fit appel à Pasteur qui avait déjà secouru les vignerons français. De même que dans ses études sur les cristaux asymétriques ou sur les différentes souches de levure, ce fut grâce à un usage intelligent de son microscope que Pasteur trouva une solution. Il découvrit que les vers à soie malades et les feuilles de mûrier dont se nourrissaient les vers étaient contaminés par des micro-organismes ; il conseilla alors de détruire tous les vers et toutes les feuilles atteints par l'infection et de reprendre l'élevage avec ce qui subsistait de vers et de feuilles non infectées. Cette vigoureuse intervention fut couronnée de succès.

Pasteur fit plus avec ses recherches que simplement faire revivre l'industrie des vers à soie. Généralisant ses conclusions, il énonça la *théorie bactérienne des maladies* – sans aucun doute la plus grande de toutes les découvertes médicales (les chimistes comme moi-même se plaisent à faire remarquer qu'il était chimiste lui aussi, et non pas médecin).

Avant Pasteur, ce que les médecins pouvaient faire de mieux pour leurs malades était de recommander du repos, une bonne nourriture, du grand air, un environnement propre, et de temps en temps, quelques soins de première urgence ; toutes choses qui étaient déjà conseillées en l'an 400 av. J.-C. par le médecin grec Hippocrate (« le père de la médecine »). Ce fut Hippocrate qui introduisit une vision rationnelle de la médecine ; il se refusait à donner des explications magiques telles que la colère d'Apollon ou la possession par les démons, et affirmait que toutes les maladies, même l'épilepsie, surnommée la « maladie sacrée », étaient causées non par une

influence surnaturelle, mais par des désordres physiques pouvant être traités en tant que tels. Cette leçon resta valable pour les générations qui suivirent.

Toutefois, la médecine fit fort peu de progrès pendant les deux mille ans qui suivirent. Les médecins se contentaient d'inciser des furoncles, de réduire des fractures et de prescrire quelques remèdes spécifiques, souvent issus de la sagesse populaire, par exemple la quinine, extraite de l'écorce de quinquina (cette écorce que les Indiens du Pérou mâchaient pour se guérir de la malaria) ou la digitale (un vieux remède de bonne femme pour stimuler le cœur). Mis à part ces quelques traitements (et le vaccin contre la variole dont je parlerai plus tard), la plupart des médicaments et des traitements prodigués par les successeurs d'Hippocrate contribuaient à augmenter le taux de mortalité plus qu'à le diminuer.

L'un des progrès les plus remarquables réalisés pendant les deux cent cinquante premières années de l'ère scientifique fut l'invention en 1819 du *stéthoscope* par le médecin français René Théophile Hyacinthe Laennec. A l'origine, ce n'était guère plus qu'un tube de bois destiné à aider le médecin à entendre et interpréter les sons produits par les battements du cœur. A l'heure actuelle, après de nombreuses améliorations, cet instrument est devenu le compagnon fidèle du médecin, aussi indispensable que la calculatrice de poche pour un ingénieur.

On comprend par conséquent que jusqu'au XIXe siècle, même les pays les plus civilisés aient été périodiquement ravagés par des épidémies, dont certaines eurent de profondes répercussions historiques. C'est une épidémie de peste à Athènes qui tua Périclès pendant la guerre du Péloponnèse et entraîna l'effondrement définitif du monde grec. La chute de Rome, elle aussi, commença par des épidémies de peste qui s'abattirent sur l'Empire romain sous le règne de Marc Aurèle. Et au XIVe siècle la peste noire tua près d'un quart de la population européenne, ce qui, s'ajoutant à la poudre à canon, aboutit à détruire les structures sociales du Moyen Age.

Bien sûr, les épidémies ne cessèrent pas lorsque Pasteur découvrit que les maladies infectieuses étaient provoquées et transmises par des micro-organismes. On continue à voir sévir des maladies endémiques, comme par exemple le choléra en Inde, et beaucoup de pays sous-développés souffrent encore d'épidémies graves. La maladie est aussi particulièrement fréquente en temps de guerre. De plus, de nouveaux organismes virulents se développent de temps en temps et se répandent à travers le monde, comme ce fut le cas pour la grande grippe de 1918 qui tua presque quinze millions de personnes, soit plus de morts que n'importe quel autre fléau de l'histoire humaine, et encore deux fois plus que n'en avait causé la précédente guerre mondiale.

Toutefois les découvertes de Pasteur marquèrent le début d'une ère nouvelle ; le taux de mortalité en Europe et aux États-Unis commença à décroître de façon appréciable, et l'espérance de vie à augmenter régulièrement. Grâce à l'étude scientifique de la maladie et de son traitement qui commença avec Pasteur, des hommes et des femmes des régions les plus développées du monde peuvent maintenant espérer vivre en moyenne jusqu'à soixante-dix ans, alors qu'auparavant la moyenne n'était que de quarante ans dans les meilleures conditions, et même de vingt-cinq ans seulement quand les conditions étaient défavorables. Après la fin de la Seconde Guerre mondiale, l'espérance de vie augmenta sans arrêt même dans les régions du monde les moins avancées.

L'IDENTIFICATION DE LA SPÉCIFICITÉ DES BACTÉRIES

En 1865, avant même que Pasteur ait énoncé sa théorie des microbes, un médecin viennois nommé Ignaz Philipp Semmelweiss avait mis au point la première stratégie antibactérienne sans savoir bien sûr ce contre quoi il se battait. Il travaillait dans la salle de maternité de l'un des hôpitaux de Vienne, où plus de 12 % des jeunes mères mouraient d'une maladie appelée *fièvre puerpérale* (ou fièvre contractée à la suite d'un accouchement). Semmelweiss avait observé avec un certain étonnement que les femmes qui accouchaient à la maison, aidées seulement par d'ignorantes sages-femmes, ne contractaient pratiquement jamais cette maladie. L'hypothèse qui lui vint à l'esprit fut confirmée par la mort d'un docteur de l'hôpital présentant tous les symptômes de la fièvre puerpérale, après qu'il se fut coupé en disséquant un cadavre. Peut-être les médecins et les étudiants qui sortaient des salles de dissection transmettaient-ils en quelque façon cette maladie aux femmes dont ils surveillaient l'accouchement ? Semmelweiss insista pour que les médecins se lavent les mains dans une solution de chlorure de chaux ; et dans l'année, le taux de mortalité dans la salle de maternité chuta de 12 % à 1,5 %.

Mais ses collègues plus âgés, au lieu de se réjouir, se retournèrent contre cet homme qui les traitait d'assassins et, humiliés par ces trop fréquents lavages de mains, expulsèrent Semmelweiss de l'hôpital (aidés en cela par ses origines hongroises, car la Hongrie était en rébellion contre l'Autriche). Semmelweiss quitta Vienne pour Budapest, où il contribua à réduire le taux de mortalité maternelle ; cependant qu'à Vienne les hôpitaux retournaient à leurs sinistres habitudes pour quelques dizaines d'années de plus. Puis, en 1865, à l'âge de quarante-sept ans, Semmelweiss mourut lui-même de septicémie à la suite d'une contamination accidentelle, quelques mois trop tôt pour voir ses hypothèses sur la transmission de cette maladie confirmées. C'est cette année-là que Pasteur découvrit des micro-organismes dans les vers à soie malades, et qu'un chirurgien anglais du nom de Joseph Lister (le fils de l'inventeur du microscope achromatique) commença lui aussi à lutter contre les microbes à l'aide de la chimie.

Lister eut d'abord recours à une substance corrosive, le *phénol* (ou acide phénique) dont il fit usage pour panser un patient souffrant d'une fracture complexe. Jusque-là, les blessures graves conduisaient toujours à des infections. Bien sûr, le phénol de Lister risquait d'abîmer les tissus entourant la blessure, mais il tuait aussi les bactéries. Ce patient eut donc une convalescence sans complications.

Lister continua sur ce succès, en adoptant la pratique de vaporiser du phénol dans la salle d'opération ; ce devait être bien désagréable pour les gens qui le respiraient, mais ceci servit à sauver quelques vies de plus. Lister, comme Semmelweiss, rencontra de l'opposition, mais l'idée de l'antisepsie était mieux acceptée à la suite des expériences de Pasteur, ce qui fait que Lister continua de gagner sa vie sans problèmes.

Pasteur avait lui-même quelques difficultés en France (il ne bénéficiait pas du statut de médecin qui protégeait Lister), mais il persuada tout de même les chirurgiens d'ébouillanter leurs instruments et de stériliser leurs bandages à la vapeur ; cette stérilisation à la vapeur remplaça avantageusement le phénol de Lister. On découvrit par la suite des antiseptiques plus doux, qui tuaient les bactéries sans abîmer inutilement les tissus. Le médecin français Casimir Joseph Davaine décrivit les propriétés anti-

septiques de l'iode en 1873, et la *teinture d'iode* (de l'iode dissous dans un mélange d'alcool et d'eau) devint d'usage courant dans les maisons. On désinfecte à présent la moindre égratignure avec ce genre de produit, ce qui sans aucun doute évite de très nombreuses infections.

En fait, la recherche d'une protection contre l'infection microbienne s'orienta de plus en plus dans le sens de la décontamination des objets ou des locaux utilisés (*asepsie*), plutôt que la désinfection de la blessure (antisepsie). En 1890, le chirugien américain William Stewart Halstead prit l'habitude d'utiliser des gants stériles pendant les opérations et en 1900, le médecin anglais William Hunter y ajouta un masque stérile pour protéger le patient contre les microbes contenus dans la respiration du médecin.

Entre-temps, le médecin allemand Robert Koch s'était mis à identifier les bactéries à l'origine de diverses maladies, en améliorant de façon significative la nature des *milieux de culture* – la source de nourriture dans laquelle on fait pousser les bactéries. Il substitua au milieu liquide utilisé par Pasteur un milieu solide, la gélatine, dans lequel on pouvait implanter les échantillons isolément les uns des autres (cette gélatine fut plus tard remplacée par de l'*agar*, une substance analogue à la gélatine extraite des algues). En déposant une seule bactérie avec une aiguille fine en un point de ce milieu, on peut obtenir une colonie homogène autour du point parce qu'à la surface solide de l'agar les bactéries ne peuvent ni bouger ni s'éloigner de leur parent d'origine, comme elles le feraient en milieu liquide. L'un de ses assistants, Julius Richard Petri, introduisit en plus l'usage de boîtes en verre peu profondes munies de couvercles pour protéger les cultures de la contamination par les spores bactériennes flottant dans l'air ; ces *boîtes de Petri* sont encore utilisées de nos jours.

Ainsi, une bactérie donnait naissance à une colonie pouvant être cultivée séparément et testée pour savoir si elle induisait une maladie dans un animal de laboratoire. Cette technique permettait non seulement d'identifier les injections, mais aussi d'expérimenter les divers moyens de tuer les espèces bactériennes qui en étaient responsables. C'est grâce à elle que Koch isola le bacille du charbon, puis en 1884 celui du choléra. La voie qu'il avait ouverte fut suivie par d'autres chercheurs, comme le pathologiste allemand Edwin Klebs qui isola l'agent infectieux de la diphtérie. Koch reçut le prix Nobel de médecine et de physiologie en 1905.

Agents chimiothérapiques

L'identité des bactéries pathogènes ayant été établie, il devenait possible de rechercher des médicaments qui détruisent ces bactéries sans mettre en danger la vie du patient. Un collaborateur de Koch, le médecin et bactériologiste allemand Paul Ehrlich, entreprit ce travail qu'il imaginait comme la recherche d'une arme douée du pouvoir magique de frapper sélectivement la bactérie et non le corps.

Ehrlich s'intéressait particulièrement aux travaux sur les colorants fixés par les bactéries, domaine de recherche essentiel à l'étude des cellules, parce que celles-ci sont incolores et transparentes et qu'on ne peut les observer en détail. Bien que des colorants cellulaires aient déjà été utilisés

par les premiers microscopistes, ils ne devinrent d'un usage facile qu'après la découverte de Perkin sur les colorants aniliques (voir chapitre 11). Ehrlich, sans être l'inventeur de cette technique, avait néanmoins contribué à l'améliorer entre 1870 et 1880, ouvrant ainsi la voie aux études de Flemming sur la mitose et à celles de Feulgen sur l'ADN chromosomique (voir chapitre 13).

En réalité, Ehrlich comptait sur le pouvoir bactéricide de ces colorants. Il pensait en effet qu'un colorant ayant plus d'affinité pour les bactéries que pour les autres cellules était susceptible de devenir létal pour ces bactéries, et qu'on pouvait même envisager de soigner un patient par injection de colorant dans le sang à une concentration suffisamment basse pour ne pas endommager ses cellules.

C'est ainsi qu'Ehrlich découvrit vers 1907 un colorant appelé le *rouge trypan*, qui présentait de l'affinité pour les trypanosomes, des micro-organismes responsables de la terrible maladie du sommeil transmise par les mouches tsé-tsé en Afrique. Injecté dans le sang à des doses convenables, le rouge trypan devenait létal pour les trypanosomes mais pas pour le patient. Toutefois, Ehrlich voulait obtenir un taux de létalité des trypanosomes encore plus élevé. Sachant que la partie toxique de la molécule de rouge trypan était sa partie azo, c'est-à-dire une paire d'atomes d'azote ($-N \equiv N-$), il étudia l'effet d'une combinaison semblable d'atomes d'arsenic ($-As \equiv As-$), l'arsenic étant chimiquement proche de l'azote mais beaucoup plus toxique. Il commença à tester au hasard des dérivés chimiques de l'arsenic en les numérotant méthodiquement au fur et à mesure. C'est en 1909 qu'un étudiant japonais d'Ehrlich, Sahachiro Hata, après avoir testé sans succès la réaction du composé numéro 606 sur les trypanosomes, obtint un effet létal de ce composé sur l'agent microbien de la syphilis (appelé *spirochète* à cause de sa forme en spirale).

Ehrlich comprit que sa découverte était bien plus importante que le remède cherché contre la trypanosomiase. La syphilis avait été le fléau caché de l'Europe pendant plus de quatre cents ans, depuis l'époque de Christophe Colomb (on soupçonnait son équipage de l'avoir contractée au contact des Indiens des Caraïbes). De plus, un silence pudique entourait cette maladie incurable, laissant le champ à une contagion incontrôlée. Ehrlich consacra le reste de sa vie à essayer de combattre la syphilis grâce au composé 606, ou Salvarsan, comme il l'appelait. Ce composé guérissait la maladie, mais non sans risque, et Ehrlich dut se battre avec les hopitaux pour qu'un usage correct en soit fait.

La guérison par voie chimique connut un nouvel essor grâce au travail d'Ehrlich. La *pharmacologie*, c'est-à-dire l'étude de l'action sur l'organisme des produits chimiques servant de médicaments, devint une science auxiliaire importante de la médecine du XXe siècle. L'arsphénamine fut le premier remède synthétique, par opposition aux remèdes d'origine végétale comme la quinine, ou minérale comme ceux utilisés par Paracelse et ses successeurs.

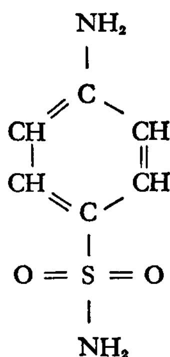
LES SULFAMIDES

Bien qu'on ait tout de suite pensé à élargir cette approche à d'autres maladies, grâce à des antidotes spécifiquement fabriqués dans chaque cas, les fabricants de remèdes nouveaux n'eurent que peu de succès au cours

des années qui suivirent la découverte d'Ehrlich. Le seul succès obtenu fut la synthèse par les chimistes allemands du *plasmochin* en 1924, et de l'*atabrine* en 1930, des substituts de la quinine contre la malaria. Ces remèdes furent utilisés pendant la Deuxième Guerre mondiale par les troupes occidentales stationnées dans les régions tropicales, lorsque l'île de Java, devenue la source mondiale de la quinine après que l'industrie de celle-ci, comme celle du caoutchouc, se fut déplacée d'Amérique du Sud en Asie du Sud-Est, fut envahie par les Japonais.

C'est en 1932 que les recherches progressèrent soudain. Le chimiste allemand Gerhard Domagk injectait divers colorants à des souris malades lorsqu'il découvrit qu'un nouveau colorant rouge, appelé le prontosil, induisait la survie de souris atteintes d'une infection mortelle due à des streptocoques hémolytiques. Il tenta alors de guérir sa propre fille qui était en train de mourir de cette même infection et y parvint. Trois ans plus tard, le pouvoir du prontosil, remède capable d'arrêter les infections de streptocoques chez l'homme, était reconnu dans le monde entier.

Bizarrement, le prontosil ne tuait pas les streptocoques dans le tube à essai, mais seulement dans le corps. Jacques Tréfouël et ses collaborateurs, de l'Institut Pasteur à Paris, en conclurent que la substance active sur les bactéries était produite par clivage du prontosil à l'intérieur du corps humain, et se mirent à isoler ce fragment actif, appelé la *sulfanilamide*. La structure chimique de ce composé, dont la synthèse avait déjà été signalée en 1908 sans attirer d'attention particulière, est reproduite ci-dessous :



Ce fut le premier d'une série de médicaments miracles. Les bactéries pathogènes furent vaincues les unes après les autres grâce à ce remède.

Les chimistes découvrirent qu'en substituant par différents groupes chimiques l'un des atomes d'H du groupe contenant le S, on obtenait une série de composés dont les propriétés antibactériennes étaient chaque fois légèrement différentes. On découvrit ainsi la *sulfapyridine* en 1937, le *sulfathiazole* en 1939, et la *sulfadiazine* en 1941. Les médecins avaient désormais à leur disposition toute une série de *médicaments sulfamidés*. Dès lors, dans les pays médicalement avancés, la mortalité due aux maladies bactériennes, en particulier dans le cas de la pneumonie pneumococcique, a diminué de façon spectaculaire.

On décerna à Domagk le prix Nobel de médecine et de physiologie en 1939. Mais il fut arrêté par la Gestapo alors qu'il écrivait sa lettre d'acceptation ; sous le régime nazi, on préférait se méfier des prix Nobel

(allez savoir pourquoi !). Quand la guerre fut finie, et qu'il fut enfin libre d'accepter cet honneur, Domagk se rendit à Stockholm où il reçut officiellement le prix.

LES ANTIBIOTIQUES

Les sulfamides n'eurent qu'une brève période de gloire, car ils furent bientôt rejetés dans l'ombre par la découverte d'agents antibactériens beaucoup plus puissants, les antibiotiques.

La matière vivante d'où qu'elle vienne (y compris celle des êtres humains) finit dans le sol où, en se dégradant et en se décomposant, elle entraîne dans les matériaux morts et leurs débris les germes des nombreuses maladies infectieuses dont étaient atteints les organismes vivants. Malgré cela très peu de microbes survivent dans le sol (à l'exception du bacille du charbon). Déjà en 1877, Pasteur avait signalé qu'il existait des souches bactériennes qui avaient un effet létal sur d'autres souches. Selon cette hypothèse, le sol devait constituer une source très variée d'organismes bactéricides, compte tenu que pour chaque hectare de sol, il y a à peu près deux mille kilos de moisissures, mille kilos de bactéries, deux cents kilos de protozoaires, cent kilos d'algues et cent kilos de levures.

Le microbiologiste franco-américain René Jules Dubos, qui s'intéressait aux substances bactéricides du sol, isola en 1939 à partir d'un micro-organisme appelé *Bacillus brevis* une substance, la *tyrothricine*, qu'il sépara en deux composants bactéricides, respectivement la *gramicidine* et la *tyrocidine*. Ceux-ci se révélèrent être des peptides contenant des acides aminés D – images en miroir des acides aminés ordinaires L qui composent la majorité des protéines naturelles.

La gramicidine et la tyrocidine furent les premiers antibiotiques produits volontairement. Cependant, douze ans auparavant, un antibiotique qui devait se révéler beaucoup plus important avait déjà été découvert ; il n'avait été cité qu'incidemment dans un article scientifique.

Le bactériologiste anglais Alexandre Fleming avait découvert un matin que ses cultures de staphylocoques (les bactéries que l'on trouve généralement dans le pus), qui étaient restées sur une paillasse, étaient contaminées par un agent bactéricide. En effet de petits cercles clairs apparaissaient dans les boîtes de culture, signe que les staphylocoques avaient été détruits. Fleming, qui s'intéressait à l'antisepsie (car il avait travaillé sur les propriétés antiseptiques d'une enzyme présente dans les larmes, le *lysozyme*), se mit tout de suite à chercher la cause de ce phénomène et découvrit qu'il était dû à une moisissure du pain, *Penicillium notatum*, celle-ci produisant une substance qu'il appela *pénicilline* et qui avait un effet létal sur les microbes. Ce résultat, publié en 1929, n'attira pas beaucoup l'attention à ce moment-là.

C'est dix ans plus tard que le biochimiste anglais Howard Walter Florey et son associé d'origine allemande, Ernst Boris Chain, s'intéressèrent à ce travail jusque-là négligé et se mirent à isoler la substance antibiotique décrite ; ils obtinrent en 1941 un extrait d'usage thérapeutique contre nombre de bactéries *Gram-positives* (bactéries réagissant positivement à une technique de coloration développée en 1884 par le bactériologiste danois Hans Christian Joachim Gram).

Du fait de la guerre, l'Angleterre ne pouvait pas produire le médicament ; Florey partit donc aux États-Unis où il aida à la mise en route d'un programme de développement des méthodes de purification et de production de la pénicilline à partir de la moisissure. Cinq cents cas de maladies furent traités à la pénicilline en 1943 ; et vers la fin de la guerre, sa production et son usage médical à grande échelle avaient commencé. Non seulement la pénicilline supplantait fort bien les sulfamides, mais elle devint (et reste encore) l'un des remèdes les plus importants de toute la pratique médicale. Son efficacité s'étend à de nombreuses maladies infectieuses, dont la pneumonie, la blennorragie, la syphilis, la fièvre puerpérale, la scarlatine, et la méningite (l'efficacité du remède s'appelle le *spectre d'action*). De plus, elle n'est pratiquement pas toxique et n'a pas d'effets secondaires indésirables (sauf chez des individus sensibles).

Fleming, Florey et Chain partagèrent le prix Nobel de médecine et de physiologie en 1945.

La découverte de la pénicilline déclencha un programme de recherche très sophistiqué pour trouver d'autres antibiotiques (ce terme fut inventé en 1942 par le bactériologiste Selman Abraham Waksman de l'université de Rutgers).

C'est Waksman qui isola en 1943, à partir d'une moisissure du sol du genre *streptomyces*, l'antibiotique appelé *streptomycine*. Cet antibiotique tue les bactéries *Gram-négatives* (celles qui ne retiennent pas la coloration de Gram) et il est d'une efficacité remarquable contre le bacille de la tuberculose, mais il est toxique et doit être utilisé avec précaution.

Waksman reçut le prix Nobel de médecine et de physiologie en 1952 pour la découverte de la streptomycine.

En 1947, un autre antibiotique, le *chloramphénicol*, fut isolé à partir des moisissures du genre *streptomyces*. Le chloramphénicol agit contre certains organismes plus petits – en particulier ceux qui provoquent la fièvre typhoïde et la psittacose (maladie infectieuse transmise à l'homme par certains oiseaux), mais sa toxicité exige de faire attention à son usage.

Par la suite, toute une série d'antibiotiques à *large spectre d'action* furent découverts par l'analyse laborieuse de quelques milliers d'échantillons du sol – l'auréomycine, la terramycine, l'actinomycine et ainsi de suite. Le premier de cette série, l'auréomycine, fut isolé par Benjamin Minge Duggar et ses collaborateurs en 1944, et mis sur le marché en 1948. Ces antibiotiques sont appelés des *tétracyclines* parce que leur formule chimique comporte quatre cycles d'atomes de carbone accolés. Ils agissent contre beaucoup d'agents infectieux différents, et grâce à eux la fréquence des maladies infectieuses est tombée à un niveau très bas dont on peut se réjouir.

Bien sûr, même si des vies humaines sont sauvées grâce aux progrès continuels que nous faisons sur les maladies infectieuses, ces mêmes personnes continuent de courir les risques associés à des désordres métaboliques, comme par exemple le diabète, dont l'incidence a augmenté par un facteur dix pendant les quatre-vingts dernières années.

LA RÉSISTANCE AUX ANTIBIOTIQUES

L'accroissement rapide du nombre de souches bactériennes résistantes aux remèdes chimiques est un problème grave qui est malheureusement apparu peu de temps après leur mise au point. On considère par exemple

que, si en 1939 tous les cas de méningites et de pneumonies pneumococciques répondaient favorablement à l'administration de sulfamides, vingt ans plus tard, seulement la moitié des cas y répondaient. De plus, il semble que les antibiotiques perdent de leur efficacité au fur et à mesure de leur usage. Cette résistance provient non pas de ce que les bactéries apprennent à résister, mais de ce que parmi celles-ci des mutants résistants qui ne sont pas tués comme la souche « normale », croissent et se multiplient. Le processus de mutation est d'ordinaire lent, et dans le cas des eucaryotes, en particulier des organismes multicellulaires, c'est par la redistribution des gènes et des chromosomes à chaque génération que les variations ou différences de phénotypes apparaissent le plus rapidement ; dans le cas des bactéries au contraire, qui se multiplient très rapidement, c'est par mutation que les changements de phénotypes surviennent.

Même si un très faible pourcentage de mutations aboutit à une nouvelle fonction (par exemple à une enzyme capable de détruire un agent chimique donné), le nombre absolu de telles mutations reste élevé parce qu'il suffit de quelques cellules souches pour produire d'innombrables descendants.

De plus, les gènes nécessaires à la production de ces enzymes se trouvent souvent portés par des plasmides et sont transférés de bactéries à bactéries, accélérant ainsi l'acquisition de la résistance.

C'est dans les hôpitaux, où l'on emploie très fréquemment des antibiotiques et où les patients ont une résistance par définition plus faible aux infections, que les souches de bactéries résistantes sont les plus dangereuses. Récemment des souches de staphylocoques particulièrement résistantes aux antibiotiques sont apparues. Ces *staphylocoques d'hôpitaux* sont un souci majeur dans les maternités par exemple, et ont fait la une des journaux en 1961 quand la célèbre actrice de cinéma Elisabeth Taylor faillit mourir d'une attaque de pneumonie provoquée par de telles bactéries résistantes.

Heureusement, lorsqu'un antibiotique échoue, un autre peut encore détruire la souche résistante. On peut fabriquer de nouveaux antibiotiques ou des modifications synthétiques des antibiotiques plus anciens qui compensent l'apparition de nouvelles mutations. L'idéal est de trouver un antibiotique pour lequel aucun mutant résistant ne puisse apparaître, de sorte qu'aucun survivant de la bactérie ne puisse se multiplier. On a mis au point des remèdes spécifiquement dans ce but : par exemple, en 1960, un variant de la pénicilline, appelé la staphcylline, qui est partiellement synthétique, a été utilisé pour tuer les bactéries jusque-là résistantes à la pénicilline. C'est parce que sa structure chimique est inconnue des bactéries que sa molécule ne peut être coupée et son activité détruite par une enzyme comme la *pénicillinase* (découverte par Chain, dont se servent les bactéries résistantes à la pénicilline ordinaire). C'est la staphcylline qui a sauvé la vie d'Elisabeth Taylor. Malgré cela, des souches de staphylocoques résistantes aux pénicillines synthétiques sont apparues, et il est à craindre que ce manège ne dure éternellement.

Tandis que de nouveaux antibiotiques ou des dérivés semi-synthétiques de ceux-ci s'ajoutent encore dans notre lutte contre les souches résistantes, on ne peut qu'espérer que la souplesse des lois de la chimie alliée à la nécessité permettra de garder la mainmise contre la souplesse avec laquelle les bactéries pathogènes s'adaptent à cette nécessité.

LES PESTICIDES

Un problème identique d'apparition de souches résistantes est apparu dans notre lutte contre des ennemis de plus grande taille, les insectes, qui sont à la fois de redoutables concurrents dans le domaine de l'alimentation et des vecteurs de maladie. La lutte chimique moderne contre les insectes a commencé en 1939 avec la découverte, par le chimiste suisse Paul Müller, du produit appelé *dichlorodiphényltrichloroéthane,* mieux connu sous ses initiales DDT, réussite pour laquelle Müller reçut le prix Nobel de médecine et de physiologie en 1948.

Mais déjà à ce moment-là, le DDT était utilisé à une très grande échelle, et des insectes résistants, comme par exemple des mouches, avaient fait leur apparition. D'autres *insecticides* étaient nécessaires – ou plutôt d'autres *pesticides,* pour employer un terme général qui couvre les produits que l'on utilise pour détruire les rats ou les mauvaises herbes.

Par ailleurs, l'utilisation excessive de la chimie dans notre lutte contre les espèces vivantes autres que l'espèce humaine est vivement critiquée. Certains pensent par exemple que la science risque d'encourager une partie sans cesse croissante de notre population à vivre grâce aux subterfuges de la chimie ; et au cas où une déficience, même temporaire, de notre organisation technologique se produirait, un désastre pourrait avoir lieu par suite de notre moindre résistance naturelle aux maladies ou aux infections dont nous étions protégés uniquement grâce aux produits chimiques.

L'écrivain scientifique américain Rachel Louise Carson a publié en 1962, au sujet des pesticides, un livre intitulé *Silent Spring* (le printemps silencieux), qui défend avec vigueur l'idée que par un usage non discriminatoire des produits chimiques, nous détruisons des espèces inoffensives ou même utiles, en même temps que celles que nous essayons réellement de détruire. La thèse défendue par Carson est qu'une destruction irréfléchie des espèces vivantes peut amener de graves changements dans le système complexe d'interdépendance des espèces, et nuire en fin de compte à l'humanité au lieu de l'aider. On appelle *écologie* l'étude de ces interactions entre espèces, et grâce au livre de Carson cette branche de la biologie a reçu une attention nouvelle.

Bien sûr, il ne s'agit pas pour autant de renoncer à nos efforts technologiques pour limiter le développement des insectes (car le prix en maladies et en famines serait trop élevé), mais de trouver des méthodes de lutte plus spécifiques et qui portent moins préjudice à la structure écologique d'ensemble.

On peut imaginer par exemple d'encourager les ennemis naturels des insectes, qu'ils soient parasites ou prédateurs, ou bien d'utiliser des sons ou des odeurs qui repoussent ou incitent au suicide les insectes, ou encore de les stériliser à l'aide de radiations. Mais de toute façon tous les efforts doivent converger vers l'insecte nuisible et vers lui seul.

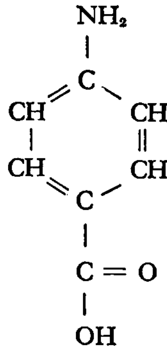
Une stratégie prometteuse, imaginée par la biologiste américaine Carroll Milton Williams, consiste à utiliser des hormones des insectes. En effet, ceux-ci passent par des mues périodiques qui comportent deux ou trois stades bien définis : la larve, la pupe et l'adulte. Ces transitions complexes sont contrôlées par des hormones, comme par exemple l'*hormone* dite *juvénile,* qui empêche la formation du stade adulte jusqu'à un moment approprié. En traitant les insectes avec cette hormone produite artificielle-

ment, on peut retarder le stade jusqu'au moment où l'insecte meurt. Comme chaque espèce d'insecte possède sa propre hormone juvénile et n'est sensible qu'à elle, une hormone juvénile donnée peut servir à attaquer une espèce précise d'insecte sans nuire aux autres organismes. Et les biologistes peuvent même fabriquer, en s'aidant de la structure de l'hormone, des substituts synthétiques à bon marché et tout aussi efficaces.

Bref, la réponse au problème posé par les effets secondaires délétères du progrès scientifique n'est pas d'abandonner ce progrès, mais de lui substituer un progrès plus subtil, appliqué intelligemment et avec précaution.

MÉCANISME D'ACTION DES AGENTS CHIMIOTHÉRAPIQUES

Le mécanisme d'action des agents chimiques antibactériens peut être résumé en disant que chaque médicament inhibe de façon compétitive une enzyme clé du micro-organisme pathogène. Le meilleur exemple d'un tel mécanisme est celui des sulfamides. Leur structure est très proche de celle de l'*acide para-aminobenzoïque* (orthographié p-aminobenzoïque), qui est représenté ci-dessous :



L'acide p-aminobenzoïque est nécessaire à la synthèse de l'*acide folique*, une substance clé dans le métabolisme des bactéries et des cellules animales. Lorsque les bactéries se servent d'une molécule de sulfanilamide au lieu d'acide p-aminobenzoïque, elles ne peuvent plus produire d'acide folique parce que l'enzyme nécessaire à ce procédé est mise hors d'usage, et en conséquence elles cessent de se multiplier. Les cellules du patient, elles, ne sont pas affectées parce que l'acide folique qu'elles utilisent provient de la nourriture et n'a pas besoin d'être synthétisé, et qu'il n'y a pas d'enzyme dans les cellules humaines pouvant être inhibée de cette façon par des concentrations faibles de sulfanilamide. Même quand la bactérie et la cellule humaine possèdent les mêmes enzymes, il existe des moyens d'attaquer la bactérie sélectivement : ou bien l'enzyme bactérienne est plus sensible à un remède donné que l'enzyme humaine, et une certaine dose tue la bactérie sans affecter les cellules humaines ; ou bien un remède correctement mis au point peut traverser la membrane de la cellule bactérienne mais pas celle de la cellule humaine. Dans le cas de la pénicilline par exemple, il y a interférence avec la construction de la paroi cellulaire spécifique de la bactérie, paroi que ne possèdent pas les cellules animales.

Pour les autres antibiotiques, un mécanisme d'action par inhibition compétitive des enzymes n'est pas absolument démontré, mais tout de même assez probable pour certains d'entre eux.

Par exemple, la gramicidine et la tyrocidine dont j'ai déjà parlé contiennent des acides aminés D atypiques qui pourraient gêner les enzymes synthétisant des produits à partir des acides aminés naturels L. La bacitracine, un autre antibiotique à structure peptidique, contient de l'ornithine, dont la structure chimique est voisine de l'arginine, et qui pourrait entrer en compétition avec celle-ci. La molécule de streptomycine, elle, contient une variété de sucre peu courante qui interférerait avec l'enzyme agissant sur l'un des sucres normaux des cellules vivantes. Le chloramphénicol, lui aussi, ressemble à l'acide aminé phénylalanine ; et enfin, une partie de la molécule de pénicilline ressemble à l'acide aminé cystéine. Dans tous ces cas, un mécanisme d'inhibition compétitive est probable.

C'est dans le cas de la *puromycine*, une substance produite par une moisissure du genre *streptomyces*, qu'un tel mécanisme a été le mieux démontré. Michael Yarmolinsky et ses collaborateurs à l'université Johns Hopkins ont pu montrer que ce composé, dont la structure ressemble beaucoup à celle des nucléotides (les éléments de base des acides nucléiques), interfère avec la synthèse des protéines en entrant en compétition avec les ARN de transfert. C'est aussi avec les ARN de transfert que la streptomycine interfère, en les induisant à lire le code génétique de façon erronée et à fabriquer des protéines sans usage. Malheureusement, cette forme d'interférence est toxique pour les cellules non bactériennes, parce qu'elle inhibe la synthèse normale des protéines indispensables. C'est pourquoi la puromycine, et la streptomycine à un moindre degré, sont trop dangereuses pour un emploi thérapeutique.

LES BACTÉRIES BÉNÉFIQUES

Il est normal que l'attention des êtres humains se concentre sur les bactéries pathogènes qui leur sont nuisibles. Toutefois, celles-ci ne représentent qu'une faible partie du nombre total de bactéries. On a calculé que pour une bactérie pathogène, il y a 30 000 bactéries non pathogènes, utiles ou même nécessaires et que, du point de vue des espèces, sur 1 400 espèces identifiées de bactéries, on n'en trouve que 150 qui sont pathogènes chez l'homme, les plantes cultivées ou les animaux domestiques.

Considérons d'abord qu'une multitude d'organismes vivants meurt à chaque instant, et que seulement un faible pourcentage de ceux-ci sont des proies pour les autres espèces animales ; en particulier, moins de 10 % des feuilles mortes et moins de 1 % du bois mort sont mangés par les animaux, et leur dégradation est en fait due à des champignons et à des bactéries. Si ces bactéries, communément appelées *bactéries saprophytes*, n'existaient pas, le monde de la vie croulerait sous le poids des débris non décomposables de plus en plus nombreux, dans lesquels une fraction également croissante des éléments vitaux s'accumulerait, et à la longue la vie elle-même s'arrêterait.

La cellulose, par exemple, n'est pas assimilable par les animaux multicellulaires, bien qu'elle soit le constituant le plus fréquent des cellules

vivantes ; même si certains animaux, comme le bétail ou les termites, vivent d'une nourriture riche en cellulose comme l'herbe ou le bois, ils n'en sont capables qu'à cause des nombreuses bactéries qui se multiplient dans leur tube digestif. Ces bactéries, qui décomposent la cellulose, lui redonnent un rôle actif dans le cycle général de la vie.

Le cycle de l'azote exige lui aussi des bactéries spécifiques. L'azote est un élément indispensable à toute vie végétale, parce qu'il sert à fabriquer des acides aminés et des protéines. Le règne animal l'obtient, déjà inclus dans les acides aminés et les protéines du règne végétal, et celui-ci l'obtient des nitrates du sol. Ces nitrates sont des sels inorganiques solubles dans l'eau qui devraient théoriquement être emportés par les eaux de pluie, ce qui rendrait la terre infertile, éliminerait la vie végétale du moins sur terre, et ne laisserait subsister que les animaux qui se nourrissent de vie aquatique.

Puisque les nitrates sont toujours présents dans le sol en dépit de l'action de millions d'années de pluie, c'est qu'il existe dans le sol une source de nitrates. Or l'azote de l'air, qui constitue une vaste réserve d'azote sous forme gazeuse, tout à fait inerte chimiquement, ne peut pas être utilisé par les plantes et les animaux pour être fixé sous des formes organiques. Ce sont en réalité des *bactéries fixatrices d'azote* qui peuvent, elles, convertir l'azote atmosphérique en ammoniac, et ensuite des *bactéries nitrifiantes* qui transforment aisément cet ammoniac en nitrates. C'est grâce à l'activité de ces diverses bactéries (et des algues bleues) que la vie sur la terre ferme est possible.

A présent, grâce à la technologie moderne – voir à ce sujet le procédé de Haber, décrit dans le chapitre 2 – les êtres humains peuvent se passer des bactéries pour fixer l'azote de l'air ; mais ce n'est qu'après que la vie terrestre a existé pendant des centaines de millions d'années. Qui plus est, la fixation industrielle de l'azote est devenue si courante que l'on pourrait craindre que les procédés naturels de dénitrification – c'est-à-dire de reconversion des nitrates en azote gazeux par d'autres bactéries encore – ne puissent fonctionner à la vitesse requise. L'accumulation des nitrates dans les rivières et dans les lacs favorise la croissance des algues et la mort d'organismes supérieurs, comme les poissons, au détriment en fin de compte d'un système écologique équilibré.

Un autre aspect bénéfique des micro-organismes (et donc des bactéries) est leur réutilisation directe à des fins alimentaires par les êtres humains, réutilisation qui date de la préhistoire. Par exemple, diverses espèces de levure sont utilisées pour la fermentation des fruits et des céréales en vin et en bière depuis l'Antiquité. Ces levures (qui sont des champignons unicellulaires eucaryotes) ont la propriété de convertir les sucres et l'amidon en alcool et en gaz carbonique. Cette même production de gaz carbonique est aussi utilisée pour faire lever la farine de blé dans la confection du pain et des pâtisseries.

La transformation du lait en yoghourt ou la fabrication de nombreux fromages sont aussi des processus de fermentation qui impliquent des moisissures et des bactéries.

L'époque moderne a vu se développer une *microbiologie industrielle* qui s'intéresse à la mise au point de souches spécifiques de bactéries et de moisissures capables de produire des substances de valeur pharmaceutique, comme les antibiotiques, les vitamines ou les acides aminés, ou de valeur industrielle, comme l'acétone, l'alcool butylique ou l'acide citrique.

Avec l'avènement du génie génétique (mentionné dans le chapitre précédent), les bactéries et autres micro-organismes pourront peut-être accroître l'efficacité de leurs propriétés bénéfiques, comme la fixation de l'azote, ou acquérir de nouvelles capacités, comme par exemple le pouvoir d'oxyder des molécules d'hydrocarbures dans certaines conditions pour dégrader les déchets du pétrole, ou encore permettre de produire industriellement des substances médicales (fractions sanguines ou hormones).

Les virus

Il peut paraître étonnant que les remèdes miracles dont nous avons parlé précédemment aient eu tant de succès contre les maladies bactériennes et si peu contre les maladies virales. Pourquoi ne serait-il pas possible de bloquer les mécanismes de reproduction des virus, comme on le fait pour les bactéries, puisqu'il faut que ces mécanismes fonctionnent pour provoquer des maladies ? La réponse est évidente une fois qu'on connaît le mode de reproduction des virus ; ce sont des parasites complets incapables de se multiplier ailleurs que dans une cellule vivante, qui n'ont que très peu et même pour ainsi dire pas du tout de mécanisme métabolique caractéristique ; et leur reproduction est totalement dépendante des matériaux fournis par la cellule qu'ils parasitent – chose qu'ils font avec une très grande efficacité, puisqu'ils peuvent se multiplier par un facteur de deux cents fois en vingt-cinq minutes. C'est pourquoi il est difficile de les priver des matériaux dont ils ont besoin ou de bloquer leur métabolisme sans affecter les cellules elles-mêmes.

Ce n'est qu'après avoir découvert des formes de vie de plus en plus simples que les biologistes ont découvert, relativement récemment, l'existence des virus. Citons d'abord, pour introduire le sujet, la découverte de l'agent de la malaria.

LES MALADIES NON BACTÉRIENNES

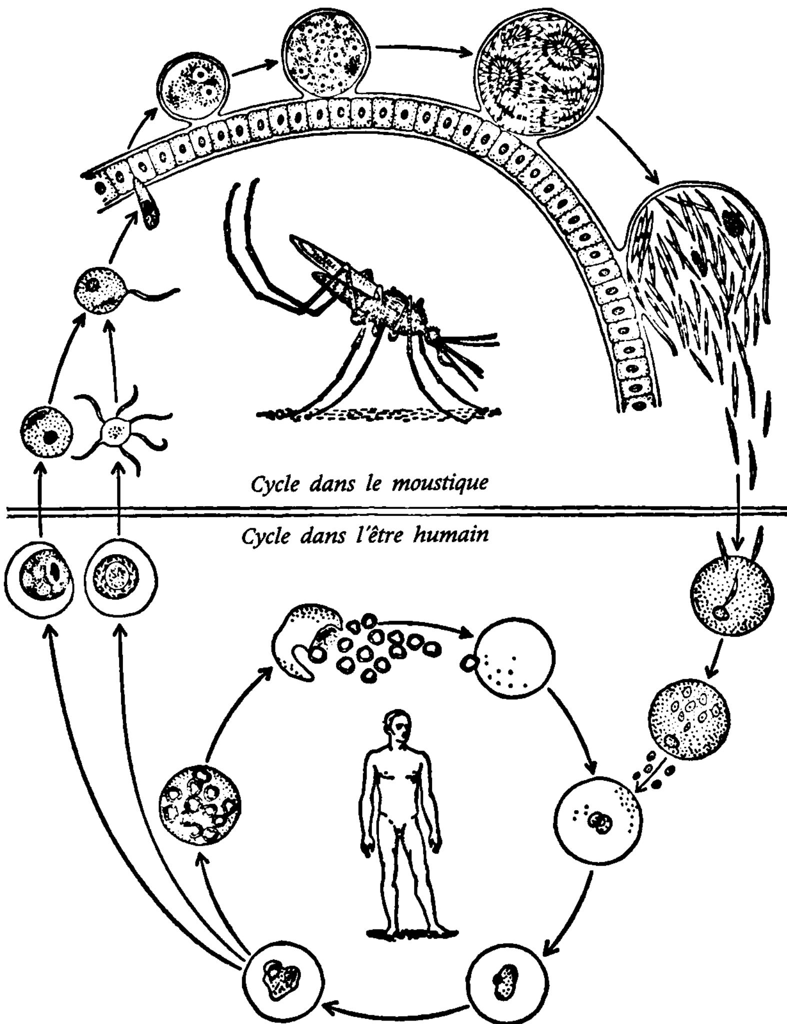
La malaria tue plus de personnes dans le monde en moyenne par an que n'importe quelle autre maladie infectieuse : selon les dernières évaluations, environ 10 % de la population mondiale souffre de cette maladie, ce qui fait trois millions de morts par an. On pensait jusqu'en 1880 que la malaria était provoquée par le « mauvais air » des régions marécageuses (*mala aria* en italien), jusqu'à ce que le bactériologiste français Charles Louis Alphonse Laveran découvre des protozoaires parasites du genre *plasmodium* dans les globules rouges des malades qui en étaient atteints (découverte pour laquelle il reçut le prix Nobel de médecine et de physiologie en 1907).

D'autre part, en 1894, un médecin anglais nommé Patrick Manson, responsable d'un hôpital missionnaire à Hong-Kong, suggéra que les moustiques qui se multiplient dans les régions marécageuses en raison de l'air humide qui y stagne interviennent peut-être dans la propagation des épidémies de malaria. Poursuivant cette idée, un médecin anglais résidant

en Inde, Ronald Ross, démontra en 1897 qu'une partie du cycle de vie du parasite agent de la malaria a lieu dans des moustiques du genre *anophèle* (voir figure 14.2), ces moustiques étant infectés par les parasites au moment où ils sucent le sang d'une personne infectieuse, le passant ensuite à toutes les personnes qu'ils piquent.

Ross reçut le prix Nobel de médecine et de physiologie en 1902 pour avoir été le premier à démontrer la transmission d'une maladie par l'intermédiaire d'un insecte jouant le rôle de *vecteur*. Ce fut une découverte très importante pour la médecine moderne, car elle prouva qu'on peut venir à bout d'une maladie en détruisant cet insecte-vecteur. On pouvait

Figure 14.2. Le cycle de vie du micro-organisme de la malaria.



désormais enrayer une épidémie de malaria en asséchant les marais où se reproduisent les moustiques, en évitant que l'eau stagne, et en tuant ces moustiques avec des insecticides. C'est ainsi qu'on a pu débarrasser de vastes parties du globe de la malaria après la Seconde Guerre mondiale, et faire diminuer le nombre de morts dues à cette maladie d'un tiers par rapport à son taux initial.

La malaria a été la première maladie infectieuse pour laquelle un agent non bactérien (dans ce cas un protozoaire) a été mis en évidence ; tout de suite après cette découverte, une autre maladie non bactérienne a été combattue de la même façon. Il s'agit de la terrible fièvre jaune qui tuait environ 95 % de ceux qu'elle atteignait, comme par exemple en 1898 pendant une épidémie à Rio de Janeiro. Lorsqu'une autre épidémie de fièvre jaune se déclara à Cuba en 1899, une commission d'enquête américaine dirigée par le bactériologiste Walter Reed fut envoyée sur place pour enquêter sur les causes de la maladie.

Reed suspectait l'existence d'un moustique vecteur, par analogie avec ce qu'on venait de découvrir pour la malaria, et son premier soin fut de démontrer que la maladie n'était pas transmise par contact direct entre les patients et les médecins, ou par l'intermédiaire des vêtements ou de la literie des malades. Puis certains des médecins se firent volontairement piquer par des moustiques qui avaient préalablement piqué un patient atteint de fièvre jaune ; ils tombèrent malades et l'un de ces courageux chercheurs, Jesse William Lazear, en mourut. Mais le coupable était identifié ; c'était le moustique *Aedes aegypti.* On put arrêter l'épidémie à Cuba. Et de nos jours, la fièvre jaune n'est plus une maladie grave dans les pays développés. L'agent de cette maladie n'est pas bactérien, mais n'est pas non plus un protozoaire, et il est de taille inférieure à une bactérie.

Le troisième exemple de maladie non bactérienne à citer est la fièvre typhoïde. Cette infection est endémique en Afrique du Nord et a atteint l'Europe via l'Espagne, au cours des longues luttes entre les Espagnols et les Maures. Cette maladie très contagieuse est un véritable fléau qui a dévasté des pays entiers. Par exemple, pendant la Première Guerre mondiale, c'est le typhus qui a chassé les armées autrichiennes de la Serbie, et non l'armée serbe elle-même ; et pendant cette même guerre et les années qui suivirent, les ravages du typhus en Pologne et en Russie aboutirent à l'anéantissement de ces pays avec autant d'efficacité qu'une action militaire (environ trois millions de morts).

C'est au début du XXe siècle que le bactériologiste français Charles Nicolle, qui dirigeait l'institut Pasteur de Tunis, constata que le typhus qui sévissait dans la ville ne provoquait pas d'épidémie dans l'hôpital. Alors que les médecins et les infirmières étaient quotidiennement en contact avec de nombreux malades atteints de typhus, il n'y avait pas de contagion à l'intérieur de l'hôpital. C'est en se demandant ce qui advenait de particulier à un malade à son entrée à l'hôpital que Nicolle se rendit compte que le changement le plus significatif était de l'ordre de l'hygiène ; en effet, le patient était soigneusement lavé et on lui enlevait tous ses habits, même et surtout si ceux-ci contenaient des poux. Nicolle émit alors l'hypothèse que le pou était le vecteur du typhus et prouva son assertion expérimentalement. Il obtint pour cette découverte le prix Nobel de médecine et de physiologie en 1928. Grâce à ses résultats et à la découverte du DDT, la fièvre typhoïde ne recommença pas les mêmes ravages pendant la Deuxième Guerre mondiale. Pour la première fois dans l'histoire, en

janvier 1944, on put enrayer une épidémie d'hiver de typhus à Naples, où l'on utilisa du DDT pour exterminer massivement les poux de la population (l'hiver, on porte davantage de vêtements et ils sont moins propres, ce qui favorise considérablement le développement des poux). Vers la fin de 1945, une autre épidémie de typhus fut stoppée de la même façon après l'occupation américaine au Japon. Ainsi, la Deuxième Guerre mondiale s'est fait remarquer dans l'histoire des guerres pour son mérite contestable d'avoir tué moins de personnes par maladies qu'à coups de fusils et de bombes.

L'agent du typhus, comme celui de la fièvre jaune, est de taille inférieure à une bactérie, ce qui nous amène à entrer maintenant dans le monde étrange et merveilleux des organismes plus petits que les bactéries.

MICRO-ORGANISMES DE TAILLE INFÉRIEURE AUX BACTÉRIES

Pour nous représenter les dimensions des objets vivants qui nous entourent, nous pouvons les considérer par ordre de taille décroissante. L'ovule humain d'un diamètre environ de cent microns (cent fois un millionième de mètre) est à peine visible à l'œil nu. Un grand protozoaire comme la paramécie est à peu près de la même taille et peut être observé en train de nager dans une goutte d'eau en microscopie optique. La cellule humaine ordinaire, environ dix fois plus petite (dix microns de diamètre), est presque invisible sans microscope. Les globules rouges sont encore plus petits – environ sept microns de diamètre au maximum. Quant aux bactéries, leur taille va de la taille moyenne d'une cellule pour les espèces les plus grandes jusqu'à des formes à peine visibles au microscope ordinaire pour les espèces les plus petites ; une bactérie en bâtonnet typique a une longueur d'environ deux microns, tandis que les toutes petites bactéries sphériques ont un diamètre à peine supérieur à quatre dixièmes de microns.

La taille de ces organismes est si petite qu'elle semble avoir atteint le plus petit volume possible permettant de faire fonctionner les mécanismes métaboliques indispensables à une vie cellulaire indépendante. Tout organisme de taille inférieure ne peut jouer le rôle d'une cellule, dont le développement est autonome, et doit donc vivre en parasite ; en effet, il lui faut se débarrasser d'une grande partie de ses mécanismes enzymatiques, comme on le ferait avec des bagages en excès par exemple, ce qui le rend alors incapable de croître et de se multiplier sur une source de nourriture artificielle, quelles qu'en soient la nature et la quantité. Un tel organisme ne peut pas être cultivé en tube à essai comme les bactéries, et ce n'est que dans une cellule vivante qui lui fournit les enzymes dont il a besoin qu'on peut le faire pousser, cette croissance ayant lieu aux dépens de la cellule hôte.

Les premières particules de taille inférieure aux bactéries ont été découvertes par un jeune pathologiste américain du nom de Howard Taylor Ricketts. En 1909, alors qu'il étudiait une maladie appelée *la fièvre pourprée des montagnes Rocheuses,* transmise par les tiques (des arthropodes hématophages qui appartiennent à la classe des arachnides), il découvrit, dans les cellules infectées de l'hôte, de très petits *organismes inclus* dans la cellule, qui sont maintenant appelés *rickettsies* en son honneur. Ricketts et ses collaborateurs découvrirent bientôt qu'une autre maladie était

provoquée par des rickettsies, le typhus. Ricketts lui-même tomba malade du typhus et mourut à l'âge de trente-neuf ans en 1910, alors qu'il était en train d'établir les preuves de cette hypothèse.

Les rickettsies (dont le diamètre varie entre deux et huit dixièmes de micron) sont de taille suffisamment grande pour être sensibles à certains antibiotiques comme le chloramphénicol et les tétracyclines, probablement parce qu'ils possèdent suffisamment de mécanismes métaboliques pour que leurs réactions aux médicaments soient différentes de celles de la cellule hôte. La thérapie par les antibiotiques a ainsi considérablement réduit le danger des maladies à rickettsies.

Les virus ont une taille qui les situe tout en bas de l'échelle des dimensions. Certains d'entre eux sont de la même taille que les rickettsies, sans qu'il y ait d'ailleurs une séparation absolue entre les rickettsies et les virus. Mais certains virus sont vraiment très petits (par exemple le virus de la fièvre jaune, dont le diamètre est de deux centièmes de micron) ; comme en moyenne leur taille n'est qu'un millième de celle d'une bactérie, ils sont bien trop petits pour qu'on puisse les détecter dans une cellule ou les voir avec un microscope optique quel qu'il soit.

Les virus n'ont pratiquement pas de mécanismes métaboliques et dépendent presque entièrement de l'équipement enzymatique de la cellule hôte et, bien que les très grands virus soient parfois sensibles à certains antibiotiques, dans la plupart des cas les médicaments n'ont pas d'effets sur eux.

On suspectait l'existence des virus depuis des dizaines d'années avant qu'on ne trouve le moyen de les voir. Pasteur, dans ses études sur la rage, ne découvrit pas d'organisme infectieux pouvant être logiquement suspecté de causer la maladie ; il suggéra que le germe dans ce cas était simplement trop petit pour être vu, ce en quoi il avait raison.

C'est en 1892, alors qu'il étudiait la *maladie de la mosaïque du tabac,* une maladie qui donne un aspect tacheté aux feuilles des plants de tabac, que le bactériologiste russe Dimitri Iosifovitch Ivanovski découvrit qu'un extrait des feuilles infectieuses pouvait transmettre la maladie par inoculation à un plant de tabac sain. Pour mettre en évidence le germe responsable, il fit passer cet extrait à travers des filtres de porcelaine capables de retenir les bactéries les plus petites, mais obtint un extrait filtré qui était encore infectieux pour les plants de tabac. Il en conclut que ses filtres devaient être déficients et perméables aux bactéries.

Le bactériologiste hollandais Martinus Willem Beijerinck répéta en 1897 cette expérience, mais sa conclusion fut que l'agent de la maladie était suffisamment petit pour passer à travers le filtre. Comme il ne pouvait rien apercevoir dans l'extrait filtré infectieux, quel que fût le microscope qu'il employait, et que lorsqu'on mettait ce liquide en culture dans un tube à essai, on ne voyait rien pousser, il émit l'hypothèse que l'agent infectieux était peut-être une petite molécule semblable à une molécule de sucre. Beijerinck appela cet agent infectieux un *virus filtrable* (*virus* étant le mot latin pour « poison »).

La même année, le bactériologiste allemand Friedrich August Johannes Löffler découvrit que la fièvre aphteuse, une maladie du bétail, pouvait être attribuée à un virus filtrable ; en 1901 Walter Reed, au cours de son travail sur la fièvre jaune, démontra que cette maladie était également due à un virus filtrable. Plus tard en 1914, le bactériologiste allemand Walther Kruse démontra que certains rhumes sont aussi dus à des virus.

Quelques années plus tard, en 1931, on connaissait environ quarante maladies qui étaient causées par des virus (dont la rougeole, les oreillons, la varicelle, la grippe, la variole, la poliomyélite et la rage), mais la nature exacte de ces virus restait mystérieuse. Finalement un bactériologiste anglais, William Joseph Elford, réussit à les retenir dans des filtres et à prouver qu'ils étaient réellement des particules matérielles. Il utilisa pour cela de fines membranes de collodion, calibrées pour retenir des particules de plus en plus petites jusqu'à ce qu'il trouve celles de ces membranes qui pouvaient retenir l'agent infectieux du liquide ; en se référant à la finesse de la membrane permettant de filtrer l'agent infectieux, il pouvait se faire une idée de la taille du virus. Grâce à ce système, il put montrer que Beijerinck avait tort, car même le virus le plus petit était de taille supérieure à la moyenne des molécules, les plus grands virus ayant à peu près la taille des rickettsies.

Après cela, pendant plusieurs années, les biologistes se querellèrent pour savoir si les virus devaient être considérés comme des particules vivantes ou non. A priori, leur capacité à se multiplier et à transmettre des maladies suggérait qu'ils étaient vivants, mais des éléments nouveaux mis en évidence par le biochimiste américain Wendell Meredith Stanley, en 1935, arguaient en faveur du contraire. L'expérience que celui-ci essayait de réaliser était d'isoler le virus de la mosaïque du tabac sous une forme aussi pure et concentrée que possible à partir d'un extrait de feuilles de tabac fortement infectées en utilisant des techniques de séparation des protéines. Or il réussit au-delà de toute espérance, car il obtint le virus sous sa forme cristallisée ! Sa préparation était aussi bien cristallisée que lorsqu'on cristallise une molécule, et malgré cela le virus restait intact, puisque, une fois redissous dans un milieu liquide, il était aussi infectieux qu'au début.

Stanley reçut pour la cristallisation de ce virus le prix Nobel de chimie en 1946 en même temps que Sumner et Northrop, d'autres cristallographes qui ont travaillé sur des enzymes (voir chapitre 12).

Malgré l'exploit de Stanley, pendant les vingt années qui suivirent, on ne sut cristalliser que des *virus végétaux* dont la structure est relativement simple (ces virus parasitent les cellules végétales). Ce n'est qu'en 1955 que le premier *virus animal*, le virus de la poliomyélite, fut cristallisé par Carlton E. Schwerdt et Frederick L. Schaffer.

Pour beaucoup de gens, y compris Stanley lui-même, la cristallisation des virus semblait la preuve que ceux-ci n'étaient composés que de protéines inertes. On n'avait jamais pu cristalliser un organisme vivant et la vie qui était par définition flexible, changeante et dynamique, semblait en contradiction avec l'aptitude à cristalliser, synonyme d'un ordre strict, rigide et fixe.

Il n'en restait pas moins que les virus étaient infectieux, qu'ils pouvaient croître et se multiplier même après avoir été cristallisés, et que la croissance et la reproduction ont toujours été considérées comme l'essence de la vie.

Il se produisit un fait nouveau en 1936 quand deux biochimistes anglais, Frederick Charles Bawden et Norman Wingate Pirie, démontrèrent que le virus de la mosaïque du tabac contenait de l'acide ribonucléique ! Fort peu, il faut dire, puisque le virus ne contenait que 6 % d'ARN pour 94 % de protéines ; mais c'était indéniablement une nucléoprotéine. Du reste, on s'aperçut bientôt que tous les virus étaient des nucléoprotéines, contenant soit de l'ARN, soit de l'ADN, soit l'un et l'autre de ces acides nucléiques.

Entre une nucléoprotéine et une protéine, il y avait presque autant de différence qu'entre la vie et la mort. Les virus contenaient donc les mêmes composants que les gènes, considérés comme l'essence de la vie ; certains grands virus ressemblaient à des chromosomes en balade pour ainsi dire, pouvant contenir jusqu'à soixante-quinze gènes contrôlant chacun la synthèse d'un constituant de structure différente (formant tantôt une fibre, tantôt un angle). La fonction ainsi que la localisation de chacun de ces gènes peuvent être déterminées à l'aide de mutations de l'acide nucléique qui rendent ce gène défectif. A l'heure actuelle, l'analyse structurelle et fonctionnelle de tous les gènes d'un virus est accessible, mais ne représente bien sûr que la première étape vers une analyse semblable des organismes à structure cellulaire, dont l'équipement génétique est beaucoup plus élaboré.

Nous pouvons nous représenter les virus dans la cellule comme des hors-la-loi qui prennent à leur compte la chimie de la cellule en se débarrassant des gènes de contrôle de celle-ci, ce qui provoque fréquemment en même temps la mort de la cellule, voire de l'organisme hôte en entier. Il arrive aussi de temps en temps que le virus remplace un gène ou une série de gènes par le sien propre, modifiant ainsi les caractéristiques de la cellule de façon héréditaire, ou qu'un virus emporte avec lui de l'ADN de la cellule bactérienne qu'il a infectée et le transporte dans la cellule suivante qu'il infecte. Ce phénomène, appelé *transduction*, a été découvert par Lederberg en 1952.

En conclusion, on peut dire que les virus sont bien vivants puisqu'ils ressemblent aux gènes qui sont porteurs des propriétés « vivantes » de la cellule. Bien sûr, cela dépend de la définition que l'on se donne de la vie, mais pour ma part, je considère qu'une molécule de nucléoprotéine capable de réplication est en vie au même titre qu'un éléphant ou un être humain.

La meilleure preuve de l'existence des virus est probablement apportée par la possibilité de les voir. On pense que la première personne à avoir vu de ses yeux un virus est un médecin écossais appelé John Brown Buist, qui publia en 1887 un compte-rendu affirmant qu'il avait réussi à voir de petits points sous le microscope dans le liquide d'une pustule de vaccination, probablement le virus de la vaccine qui figure parmi les plus grands virus que l'on connaisse.

Un microscope ordinaire ne suffisait pas pour étudier ou même simplement pour voir un virus typique ; il fallait pour cela un autre instrument, qui fut finalement mis au point vers la fin des années trente, le microscope électronique, qui permet d'obtenir des agrandissements par un facteur cent mille, et de distinguer des objets dont le diamètre est de l'ordre du nanomètre (un millième de micron).

Malgré tout, le microscope électronique a des défauts : par exemple, l'objet doit être placé sous vide et cette déshydratation inévitable peut en changer la forme ; dans le cas des cellules, il faut les découper en lamelles très fines pour pouvoir les observer ; cette image n'a que deux dimensions ; et enfin, le contraste par rapport au bruit de fond est faible parce que les électrons ont tendance à traverser le matériel biologique.

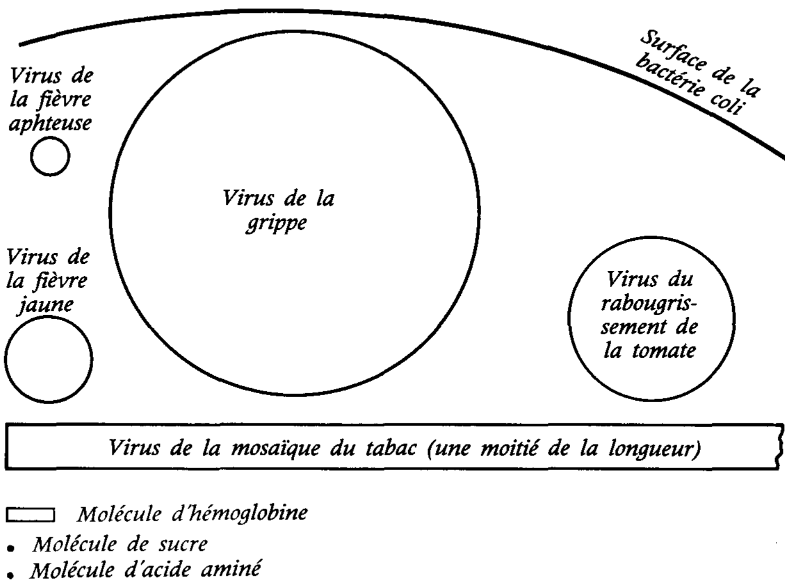
Certaines de ces difficultés furent résolues de façon ingénieuse en 1944, par l'astronome et physicien américain Robley Cook Williams et le microscopiste électronique Ralph Walter Graystone Wyckoff. Williams, qui était astronome, se dit que par analogie avec les cratères et les montagnes de la Lune, qui sont mis en relief par leur ombre quand la lumière du

Soleil les frappe obliquement, les virus pourraient être observés en trois dimensions avec un microscope électronique, si on pouvait d'une façon ou d'une autre leur faire projeter des ombres. La solution adoptée par les expérimentateurs fut de vaporiser du métal obliquement sur les particules virales déposées sur une grille, le flux de métal dessinant une surface claire – une ombre – derrière chaque particule virale ; la longueur de l'ombre indique la hauteur de la particule faisant obstacle, et l'épaisseur de la couche métallique condensée en une fine pellicule autour du virus dessine avec précision les contours des particules virales par rapport au fond.

Grâce à ces images en relief, on a pu mettre en évidence les formes de divers virus (figure 14.3). Le virus de la vaccine a la forme d'un petit tonneau et il a environ vingt-cinq centièmes de micron de longueur, ce qui fait à peu près la même taille que les rickettsies. Le virus de la mosaïque du tabac est un fin bâtonnet de vingt-huit centièmes de micron de long sur quinze millièmes de micron de large. Les plus petits virus comme la poliomyélite, la fièvre jaune et la fièvre aphteuse sont de petites sphères d'un diamètre variant entre deux centièmes et deux centièmes et demi de micron, soit bien plus petits que la taille évaluée pour un gène humain. Le poids de ces virus ne dépasse pas cent fois celui d'une molécule de protéine de taille moyenne, comme par exemple pour le virus de la mosaïque du brome, dont le poids moléculaire est de quatre millions et demi, et qui n'a que un dixième de la taille du virus de la mosaïque du tabac.

Il faut aussi citer la mise au point, en 1959, par le cytologiste finlandais Alvar P. Wilska, d'un microscope électronique faisant usage d'électrons

Figure 14.3. Tailles relatives de substances simples de protéines, de diverses particules et bactéries (38 millimètres sur cette échelle font 1/10 000 de millimètre dans la vie réelle).



à vitesse relativement lente, lesquels sont moins pénétrants que les électrons à grande vitesse et permettent de mettre en évidence certains des détails internes de la structure des virus ; en 1961, d'autre part, le cytologiste français Gaston Dupouy inventa un moyen d'observer au microscope électronique des cellules vivantes comme des bactéries en les plaçant dans des capsules remplies d'air (ces images ne sont toutefois pas détaillées en l'absence d'ombrage métallique).

Le microscope électronique usuel est un système par *transmission :* les électrons passent à travers la coupe fine à observer et sont enregistrés de l'autre côté. En utilisant un rayon électronique à faible énergie qui parcourt l'objet à étudier, comme ce serait le cas pour un téléviseur, on peut enregistrer l'émission des électrons par la surface de ce matériel ; avec ce type de microscope appelé *microscope électronique à balayage*, on peut obtenir une information détaillée sur la surface des objets. Ce procédé, suggéré par le scientifique anglais C.W. Oatley en 1948, a été utilisé à partir de 1958.

LE RÔLE DES ACIDES NUCLÉIQUES

A présent, les virologistes se sont mis à décomposer les virus et à les recomposer. Par exemple, le biochimiste américain d'origine allemande Heinz Fraenkel-Conrat de l'université de Californie, en collaboration avec Robley Williams, a découvert qu'un traitement chimique doux décompose la protéine principale du virus de la mosaïque du tabac en environ 2 200 sous-unités de 158 acides aminés chacune et de poids moléculaire 18 000 daltons, dont la séquence en acides aminés exacte et complète a été établie en 1960. Ces sous-unités remises en solution tendent à se rassembler entre elles pour reformer un long tube creux (qui est la forme sous laquelle elles existent dans le virus intact), les liaisons entre ces sous-unités étant favorisées par la présence d'atomes de calcium et de magnésium.

De façon générale, les sous-unités des protéines virales tendent à se combiner entre elles pour former des schémas géométriques, par exemple en fragments d'hélice pour le virus de la mosaïque du tabac dont je viens de parler, en un solide régulier fait de douze pentagones pour les soixante sous-unités de la protéine du virus de la poliomyélite, et en un icosaèdre à vingt faces pour les vingt sous-unités du virus de *Tipula iridescens*.

Les protéines virales sont creuses. L'hélice protéique du virus de la mosaïque du tabac, par exemple, est composée de cent trente tours de chaîne polypeptidique, abritant une longue cavité linéaire à l'intérieur de laquelle se trouve l'acide nucléique du virus (ADN ou ARN) d'une longueur fixe de 6 000 nucléotides (même si on peut détecter dans ce virus, comme l'a fait Sol Spiegelman, une molécule d'ARN longue seulement de 470 nucléotides et capable de réplication).

Fraenkel-Conrat essaya de déterminer le pouvoir infectieux de chacun des deux composants du virus de la mosaïque du tabac (acide nucléique et protéine) pris isolément, après que ces deux composants furent séparés l'un de l'autre. Il conclut en première analyse que ce pouvoir était nul lorsqu'ils étaient séparés, mais que jusqu'à 50 % de l'infectivité d'origine du virus pouvait être retrouvée lorsqu'on remélangeait protéine et acide nucléique.

Que s'était-il passé ? La protéine virale et l'acide nucléique, quasi morts lorsqu'ils étaient séparés, semblaient revenir à la vie, du moins en partie, une fois remélangés ensemble. Le grand public vit dans l'expérience de Fraenkel-Conrat la création d'un organisme vivant à partir de matériaux morts. Mais, comme nous allons le voir, il y avait méprise.

Manifestement, il y avait eu une sorte de reconstruction à partir des protéines et des acides nucléiques, chacun des deux composants ayant un rôle à jouer dans l'infection. Il restait à déterminer leurs rôles respectifs et l'importance de ces rôles.

Fraenkel-Conrat répondit à la question par une expérience très élégante. Il mélangea la partie protéique d'une souche virale avec l'acide nucléique d'une autre souche : ces deux éléments se combinèrent pour former un virus infectieux ayant des propriétés hybrides ! Ce virus était de même nature que la souche ayant fourni la protéine pour sa virulence (c'est-à-dire le degré de son pouvoir infectieux envers les plants de tabac), et identique à la souche ayant fourni l'acide nucléique pour la maladie produite (c'est-à-dire le type de mosaïque produite sur la feuille).

Ce résultat était en accord avec les idées des virologistes sur les rôles respectifs des acides nucléiques et des protéines des virus. En effet, lorsqu'un virus s'attaque à une cellule, il semble que la coque protéique ou capside lui serve à s'ancrer sur la cellule hôte et à s'ouvrir un passage dans la cellule, tandis que l'acide nucléique a le pouvoir d'envahir la cellule et de mettre en route la synthèse des particules virales nouvelles.

De plus, la nouvelle génération de virus produite par les feuilles infectées avec le virus hybride de Fraenkel-Conrat avait les caractéristiques non du virus hybride, mais seulement de la souche ayant fourni l'acide nucléique, c'est-à-dire qu'elle montrait le même degré d'infectivité et les mêmes symptômes de maladie que cette souche. En d'autres mots, c'était l'acide nucléique du virus hybride qui avait fourni l'information pour la construction de la nouvelle capside protéique, qui était la protéine de sa propre souche et non de la souche avec laquelle il avait été combiné dans l'hybride.

C'était une preuve de plus en faveur de l'idée que l'acide nucléique est bien la partie vivante des virus, et d'ailleurs, de n'importe quelle nucléoprotéine. Fraenkel-Conrat put même montrer ultérieurement que l'acide nucléique pur est capable de produire à lui tout seul une infection, quoiqu'avec peu d'efficacité (environ 0,1 % de ce que produirait le virus intact). Ceci est probablement dû à ce que de temps en temps une molécule d'acide nucléique entre d'elle-même dans la cellule sans qu'on sache comment.

Ainsi, en reconstruisant un virus à partir de son acide nucléique et de sa protéine, on ne crée pas la vie ex nihilo, puisque celle-ci est déjà présente sous la forme de la molécule d'acide nucléique. Cette dernière est protégée par la protéine virale contre l'action des enzymes hydrolysantes de l'environnement (*nucléases*), et l'efficacité du processus d'infection et de reproduction est augmentée. On pourrait comparer la molécule d'acide nucléique à un être humain et la partie protéique à une automobile ; leur combinaison facilite les voyages d'un endroit à l'autre, mais alors que l'automobile toute seule ne pourrait jamais faire le voyage, l'homme pourrait le faire à pied, ce qu'il fait parfois, quoique l'automobile l'aide beaucoup.

Les informations les plus précises et les plus détaillées sur le mécanisme d'infection d'une cellule par un virus proviennent des études sur les virus appelés bactériophages, qui ont été découverts en 1915 par le bactériologiste anglais Frederick William Twort, et indépendamment en 1917 par le bactériologiste canadien Félix Hubert d'Hérelle. Si bizarre que cela paraisse, ces virus sont des micro-organismes qui parasitent les bactéries qui sont aussi des micro-organismes. D'Hérelle les a appelés des *bactériophages,* mot dont l'étymologie grecque signifie « mangeurs de bactéries ».

Les bactériophages sont de merveilleux objets d'étude parce qu'ils peuvent être mis en culture en même temps que leur hôte dans un tube à essai. Le mécanisme de l'infection et de la multiplication de ces bactériophages est décrit ci-dessous.

Un bactériophage typique (aussi appelé *phage*) a la forme d'un petit têtard avec une tête émoussée et une sorte de queue. Grâce aux images en microscopie électronique, on sait que le bactériophage s'attache d'abord à la surface de la bactérie à l'aide de ses fibres caudales, probablement parce que la répartition des charges électriques (dues aux charges des acides aminés), à l'extrémité de la queue du phage, concorde exactement avec la répartition des charges électriques de certaines portions de la surface bactérienne ; ces répartitions de charges opposées et qui s'attirent sont si bien adaptées l'une à l'autre qu'elles provoquent une sorte de déclic analogue à ce qui se produirait lorsque les dents d'un mécanisme d'embrayage s'enclenchent. Une fois agrippé à sa victime par l'extrémité de sa queue, le phage se découpe une petite ouverture dans la paroi cellulaire, au moyen d'une enzyme qui désagrège les molécules à cet endroit. Puis, sur les images de microscopie électronique, on ne voit plus rien qui se passe : le phage, ou du moins sa coque toujours visible, reste attaché à l'extérieur de la bactérie, et aucune activité n'est visible à l'intérieur de la cellule bactérienne. Mais au bout d'une demi-heure, la cellule éclate et des centaines de virus ayant atteint leur maturité s'en échappent.

Seule la coque protéique du virus parasite est restée en dehors de la cellule, et l'acide nucléique qui se trouvait à l'intérieur de la coque protéique s'est en fait glissé dans la bactérie à travers le trou de la paroi. Le bactériologiste américain Alfred Day Hershey a démontré à l'aide de marqueurs radioactifs que le matériel envahissant la cellule est un acide nucléique dépourvu de protéines. Il commença par marquer les bactériophages avec des molécules de phosphore et de soufre radioactives, en les faisant pousser dans des bactéries qui ont incorporé ces radio-isotopes à partir de leur milieu nutritif ; or, le phosphore est incorporé à la fois dans les protéines et les acides nucléiques, tandis que le soufre n'apparaît que dans les protéines parce que les acides nucléiques ne contiennent pas de soufre. Le résultat obtenu fut que le phage marqué avec ces deux isotopes après infection de la bactérie donna une progéniture marquée au phosphore radioactif mais non au soufre radioactif, d'où la conclusion que l'acide nucléique parental avait pénétré dans la cellule mais non la protéine, et que toutes les protéines de la progéniture étaient fournies par l'hôte bactérien.

Une fois de plus, le rôle dominant des acides nucléiques dans le processus de vie était démontré. Visiblement, seul l'acide nucléique du phage avait pénétré dans la cellule, et dirigeait de là la construction des nouveaux virus – protéines ou autres matériaux – à partir du matériel cellulaire.

Il y a même des cas où l'agent infectieux n'est que de l'acide nucléique, comme par exemple pour l'agent de la maladie en fuseau de la pomme de terre dont on a démontré qu'il était un virus de toute petite taille. Le microbiologiste T.O. Diener a suggéré en 1967 que ce virus était un brin d'ARN dépourvu de protéines, et il a appelé *viroïdes* ces morceaux d'acides nucléiques possédant un pouvoir infectieux (sans l'aide d'une protéine). On connaît environ une demi-douzaine de maladies végétales qui peuvent être attribuées à des infections par des viroïdes.

Le poids moléculaire de l'un de ces viroïdes a été estimé à 130 000 (soit 1/300 de celui du virus de la mosaïque du tabac) et sa longueur à 400 nucléotides environ, des dimensions qui suffisent à permettre la réplication et donc la vie, et font probablement des viroïdes les plus petits êtres vivants qui existent.

De tels viroïdes sont peut-être impliqués dans certaines des maladies animales dites dégénératives, encore assez mal connues et provoquées par des agents infectieux à évolution « lente », c'est-à-dire qui mettent longtemps à produire des symptômes ; ceci pourrait être le résultat du faible taux d'infectivité de ces courtes chaînes d'acides nucléiques sans enveloppe.

Les réactions immunes

Les virus, qui sont nos ennemis les plus terribles parmi les êtres vivants, n'offrent presque aucune prise aux traitements par des remèdes chimiques ou autres artifices à cause de leur étroite association avec les cellules du corps humain. Et pourtant nous réussissons à nous défendre contre eux, même dans les conditions les plus défavorables, grâce aux défenses extraordinaires dont notre organisme humain est pourvu par la nature contre les maladies.

Prenons par exemple le cas de la peste noire du XIVe siècle ; elle s'abattit sur une Europe vivant dans une crasse épouvantable, dans l'ignorance de toutes les notions modernes de propreté et d'hygiène, sans égouts, sans traitements médicaux rationnels, surpeuplée, impuissante contre le fléau ; et si certains pouvaient s'enfuir des villages où régnait l'épidémie, ils ne faisaient que la répandre plus vite et plus loin lorsqu'ils étaient eux-mêmes malades. Malgré cela, les trois quarts de la population ont résisté avec succès à l'épidémie, et on s'étonne, dans ces conditions, non pas de ce qu'une personne sur quatre soit morte mais de ce que trois personnes sur quatre aient survécu.

Il existe donc probablement une sorte de résistance naturelle aux maladies. D'une part, parmi les personnes exposées à une maladie contagieuse grave, même si certaines en meurent et d'autres en sont gravement atteintes, il en est qui ne sont atteintes que de façon bénigne, ou même parfois sont complètement immunes (cette immunité peut être innée ou acquise). D'autre part, il suffit d'une seule atteinte par une maladie comme la rougeole, les oreillons ou la varicelle pour rendre immune une personne contre cette maladie, pour la vie entière.

Ces trois maladies, qui sont d'origine virale, sont des infections relativement sans gravité et rarement mortelles ; la plus dangereuse des

trois, la rougeole, ne produit que des symptômes bénins, en tout cas chez l'enfant. Comment le corps lutte-t-il contre ce virus et se constitue-t-il des défenses contre lui pour toute la vie ? La réponse à cette question est un épisode passionnant de la science médicale moderne, et il nous faut remonter à la conquête de la variole pour commencer à raconter cette histoire.

LA VARIOLE

La variole était une maladie particulièrement redoutée jusqu'à la fin du XVIII[e] siècle, parce qu'elle était souvent mortelle et que ceux qui en guérissaient étaient défigurés à vie : leur peau restait grêlée dans les cas bénins, et leurs visages perdaient toute trace de beauté ou même d'humanité dans les cas graves. Ces visages marqués par la variole étaient chose courante, et l'on vivait dans la crainte de la contagion lorsqu'on n'avait pas encore attrapé cette maladie.

En Turquie, au XVII[e] siècle, des tentatives pour se protéger contre les formes les plus graves de variole avaient déjà été faites en s'écorchant avec du sérum extrait des pustules d'une personne peu atteinte ; mais si certains attrapaient la maladie sous une forme bénigne, d'autres enduraient la défiguration et la mort qu'ils avaient précisément essayé d'éviter. C'était prendre beaucoup de risques pour échapper à la maladie et c'est une preuve de l'horreur qu'elle représentait.

Lors d'un voyage en Turquie en 1718 avec son mari, lequel était envoyé là-bas pour quelque temps comme ambassadeur d'Angleterre, une célèbre jeune femme, Lady Mary Wortley Montagu, apprit l'existence de cette pratique et fit procéder à cette inoculation sur ses propres enfants qui s'en sortirent sans dommage. Mais l'idée ne se progagea pas en Angleterre, peut-être parce que Lady Montagu était considérée comme une excentrique notoire. C'était aussi ce que les Anglais pensaient de Zabdiel Boylston, un médecin américain qui fit procéder à l'inoculation de 241 personnes pendant une épidémie de variole à Boston, mais qui fut considérablement critiqué parce que six de ces personnes moururent.

Une autre manière d'éviter la variole était pratiquée dans le Gloucestershire, où l'on pensait qu'en contractant une maladie appelée la vaccine, normalement observée chez la vache et qui se transmettait parfois aux êtres humains, on devenait immune à la variole en même temps qu'à la vaccine ; c'était d'autant plus merveilleux que cette maladie ne provoquait pratiquement pas de pustules et ne laissait aucune marque. Edward Jenner, un médecin du Gloucestershire, s'intéressa à cette pratique qui lui semblait justifiée, car il avait effectivement remarqué une fréquence plus faible des traces de la variole chez les filles de ferme particulièrement exposées à la vaccine (on peut supposer que la mode du XVIII[e] siècle consistant à célébrer la beauté des paysannes était due en partie à leur beau teint clair dans un monde marqué par la variole).

Etait-il possible que la vaccine et la variole se ressemblent suffisamment pour que les défenses fabriquées par le corps contre l'une protègent également contre l'autre ? Jenner commença à vérifier cette idée tout en prenant de grandes précautions (probablement en l'expérimentant d'abord sur sa propre famille). En 1796 il tenta l'expérience décisive. Il inocula d'abord le liquide d'une pustule de vaccine (recueilli sur la main d'une

fermière qui en était atteinte) à un petit garçon de huit ans, James Phipps. Puis, deux mois plus tard, il procéda à la partie cruciale de son expérience, en inoculant volontairement le jeune James avec la véritable variole. Le garçon n'attrapa pas la maladie car il était immune !

Jenner appela ce procédé *vaccination,* du nom latin pour « vaccine » (*vaccinia*). La vaccination se répandit à travers l'Europe comme une traînée de poudre, et fut l'un des rares cas de progrès médical à être adopté aussi facilement et immédiatement, ce qui reflète bien la peur provoquée par cette maladie et l'ardeur du public à essayer n'importe quel remède promettant la guérison. Le corps médical lui-même n'opposa qu'une faible résistance à la vaccination – mis à part quelques difficultés d'ordre symbolique, comme par exemple le refus en 1813 d'admettre Jenner au Royal College of Physicians de Londres, sous prétexte que son niveau n'égalait pas celui d'Hippocrate et de Galien.

De nos jours, grâce à la vaccination, la variole ne sévit plus, faute de trouver des hôtes non immunisés. Il n'y a pas eu un seul cas de variole aux États-Unis depuis 1949, ni d'ailleurs dans le monde depuis 1977. Des échantillons du virus ne subsistent que dans certains laboratoires, où bien sûr il peut encore y avoir des accidents.

LES VACCINS

Il s'écoula un siècle et demi pendant lequel les essais pour appliquer ce type d'inoculation à d'autres maladies graves furent infructueux. C'est Pasteur qui fit progresser ce travail en découvrant, un peu par accident, que la gravité d'une maladie pouvait diminuer si l'on affaiblissait le microbe qui la produisait.

Il travaillait alors sur la bactérie du choléra du poulet et utilisait une préparation très concentrée, dont la virulence était suffisante pour qu'une faible quantité de celle-ci, injectée sous la peau d'un poulet, le tue en un seul jour. Ayant utilisé dans un cas une culture qui avait attendu pendant une semaine, il constata alors que les poulets n'étaient que légèrement malades et qu'ils se rétablissaient ensuite. Puisque sa culture était abîmée, Pasteur prépara un autre lot virulent, mais n'obtint pas l'effet létal escompté sur les poulets, qui avaient guéri de la dose de bactéries « abîmées ». Donc, l'infection avec des bactéries affaiblies avait doté les poulets de défenses contre les bactéries encore pleinement efficaces.

En un sens, ce que Pasteur avait produit était une « vaccine » artificielle pour cette forme particulière de « variole ». Il appela son procédé *vaccination,* en reconnaissance de la dette de principe qu'il avait envers Jenner, et bien que cela n'eût rien à voir avec la vaccine. Depuis lors, le terme a été utilisé de façon très générale pour désigner les inoculations protégeant contre toutes sortes d'autres maladies, et on appelle *vaccins* les préparations employées à cette fin.

Pasteur développa encore d'autres méthodes pour affaiblir (ou atténuer) des agents de maladie. Il découvrit par exemple qu'en cultivant les bactéries du charbon à haute température, la souche était suffisamment affaiblie pour qu'on puisse s'en servir pour immuniser des animaux. Le charbon avait été jusqu'alors une maladie si contagieuse et si mortelle qu'il fallait abattre et brûler le troupeau entier aussitôt que l'une des bêtes l'avait contractée.

La plus célèbre des victoires de Pasteur fut celle remportée sur une maladie virale appelée hydrophobie ou *rage* (du mot latin signifiant « délirer », parce que cette maladie attaque le système nerveux et produit des symptômes du même genre que la folie). Le sujet mordu par un chien enragé était en proie à des symptômes violents après une période d'incubation d'un ou deux mois, et mourait presque toujours d'une mort atroce.

Pasteur se servait d'animaux vivants pour cultiver cet agent parce qu'il ne distinguait pas de microbe au microscope pouvant être l'agent de la rage (il ignorait bien sûr l'existence des virus) : il injectait donc le fluide infectieux dans le cerveau d'un lapin, le laissait incuber, broyait la moelle épinière du lapin, injectait cet extrait dans le cerveau d'un autre lapin, et ainsi de suite. Il obtint à l'aide de ce procédé des préparations atténuées du microbe qu'il faisait vieillir avant de les tester, et ceci jusqu'à ce que l'extrait ne puisse plus causer de maladie décelable chez le lapin. Il injecta ensuite une préparation de l'agent de la rage dont le pouvoir infectieux était maximal à l'un de ces lapins, et constata que l'animal était immune.

Il eut l'occasion d'essayer son traitement sur un être humain en 1885, lorsqu'on lui amena un jeune garçon âgé de neuf ans, Joseph Meister, qui avait été gravement mordu par un chien enragé. Pasteur, après bien des angoisses et des hésitations, le traita par des inoculations successives de l'agent infectieux de moins en moins atténué, dans l'espoir de lui construire une immunité avant que la période d'incubation ne se termine. Il y réussit, en tout cas le garçon survécut. Meister devint le gardien de l'Institut Pasteur, mais il mit fin à ses jours en 1940 devant l'armée nazie qui lui ordonnait d'ouvrir le tombeau de Pasteur.

En 1890, c'est sur une autre idée que se mit à travailler Emil von Behring, un collaborateur de Koch. Pourquoi prendre le risque d'injecter dans un être humain le microbe lui-même, même sous une forme atténuée ? En supposant que l'agent de la maladie provoque dans le corps la fabrication d'une substance provoquant l'immunité, ne serait-il pas aussi efficace d'extraire la substance produite par un animal après que celui-ci a été infecté avec l'agent, et de l'injecter dans le patient humain ?

Le projet de von Behring était réalisable. Il appela *antitoxine* la substance provoquant l'immunité qui se trouvait dans le sérum sanguin. Il obtint la production d'antitoxines animales contre le tétanos et la diphtérie, et son premier essai de l'antitoxine diphtérique sur un enfant atteint de cette maladie fut un succès spectaculaire ; le traitement fut immédiatement adopté, et le taux de mortalité dû à la diphtérie diminua considérablement.

Paul Ehrlich (celui qui plus tard devait découvrir l'arme magique contre la syphilis) était un collaborateur de von Behring, et c'est probablement lui qui a calculé les doses d'antitoxines à injecter. Il se brouilla avec von Behring quelque temps plus tard (c'était un individu irascible qui se brouillait facilement avec les autres) et continua de travailler tout seul sur la théorie de la thérapie sérologique. Von Behring reçut le prix Nobel de médecine et de physiologie en 1901, la première année où ce prix fut décerné. Ehrlich reçut lui aussi un prix Nobel de médecine et de physiologie en 1908, qu'il partagea avec un biologiste russe.

L'immunité conférée par une antitoxine ne dure qu'aussi longtemps que l'antitoxine reste dans le sang. Le bactériologiste français Gaston Ramon découvrit qu'en traitant directement la toxine de la diphtérie ou du tétanos avec du formaldéhyde ou une température élevée, on pouvait changer

sa structure de telle sorte que la nouvelle substance (appelée *anatoxine*) puisse être injectée sans danger à un patient humain. L'antitoxine ainsi fabriquée par le patient subsiste dans le sang plus longtemps que celle d'un animal, et d'autres doses de ce toxoïde peuvent être réinjectées pour renouveler l'immunité quand c'est nécessaire. Grâce à l'usage des anatoxines à partir de 1925, la diphtérie perdit son caractère effrayant.

Les réactions sérologiques pouvaient aussi servir à détecter la présence de la maladie. L'exemple le plus connu est celui du *test de Wasserman*, introduit par le bactériologiste allemand August von Wasserman en 1906 pour la détection de la syphilis. Ce test repose sur des techniques mises au point par le bactériologiste belge Jules Bordet, qui travaillait sur des fractions de sérum appelées *complément*, et qui s'avérèrent être un système complexe d'enzymes en interrelation les unes avec les autres. Bordet reçut pour cela le prix Nobel de médecine et de physiologie en 1919.

La lutte laborieuse de Pasteur contre le virus de la rage prouve combien il est difficile de travailler sur des virus. Si les bactéries peuvent être mises en culture, manipulées et atténuées en milieu artificiel dans un tube à essai, les virus, eux, ne peuvent pousser que dans un tissu vivant. Par exemple, dans le cas de la variole, les hôtes vivants pour le matériel à expérimenter (le virus de la vaccine) sont des vaches et des fermières, et dans le cas de la rage, ce sont des lapins que Pasteur utilisa. Mais les animaux vivants sont un moyen pour le moins malhabile, coûteux et lent de cultiver des micro-organismes.

C'est au début de ce siècle que le biologiste français Alexis Carrel a réussi l'exploit de maintenir en vie des morceaux de tissus en tube à essai – exploit qui s'est révélé d'une immense valeur pour la recherche médicale et grâce auquel ce chercheur s'est fait une réputation considérable. C'est son travail de chirurgien qui a amené Carrel à s'intéresser à cette expérience, et il a été le premier à mettre au point des méthodes de transplantation de vaisseaux sanguins et d'autres organes animaux, découverte pour laquelle il a reçu le prix Nobel de médecine et de physiologie en 1912. Pour maintenir en vie l'organe excisé avant de le transplanter, il mit au point un moyen de nourrir le tissu en le perfusant avec du sang et en lui fournissant tous les extraits et ions dont il avait besoin. Par la même occasion, il développa avec l'aide de Charles Augustus Lindbergh un *cœur mécanique* primitif capable de pomper le sang à travers les tissus.

Le dispositif de Carrel pouvait maintenir en vie un morceau de cœur d'embryon de poulet pendant trente-quatre ans – bien plus longtemps que la durée réelle de la vie d'un poulet. Carrel réussit également à faire pousser des virus dans ses cultures de tissus, mais malheureusement, les bactéries y poussaient aussi ; et pour obtenir le virus pur, les précautions d'asepsie à prendre étaient si complexes qu'il était finalement plus facile d'utiliser des animaux vivants.

Toutefois, l'idée de l'embryon de poulet était dans l'air, et on en vint bientôt à utiliser l'embryon de poulet entier comme milieu de culture au lieu de seulement un morceau de tissu. En effet, un embryon de poulet est un organisme indépendant protégé par la coquille de l'œuf et équipé de ses propres défenses naturelles contre les bactéries ; il est également bon marché et facile à obtenir en grande quantité. Le pathologiste Ernest William Goodpasture et ses collaborateurs, à l'université Vanderbilt (Nashville, Tennessee), réussirent à cultiver un virus dans un embryon

de poulet en 1931, et pour la première fois la culture des virus purs devint une pratique presque aussi facile que celle des bactéries.

C'est en 1937 qu'eut lieu la première grande victoire médicale faisant appel aux cultures de virus dans des œufs fertiles. Les bactériologistes de l'institut Rockefeller étaient encore à la recherche d'une meilleure protection contre le virus de la fièvre jaune parce que le moustique était impossible à éradiquer complètement ; les singes infectés constituaient un réservoir de virus qui menaçait constamment de répandre cette maladie dans les régions tropicales. Le bactériologiste sud-africain Max Theiler, qui travaillait dans cet institut à produire un virus atténué de la fièvre jaune, fit passer le virus à travers deux cents souris et cent embryons de poulets, et il obtint un mutant capable de ne causer que des symptômes bénins tout en conférant une immunité complète contre la fièvre jaune. Theiler reçut pour ce résultat le prix Nobel de médecine et de physiologie en 1951.

Mais finalement rien ne vaut une culture en tube à essai parce qu'elle est rapide et efficace, et que les conditions sont faciles à contrôler ; c'est pourquoi John Franklin Enders, Thomas Huckle Weller et Frederick Chapman Robbins, à la Harvard Medical School, revinrent à l'approche de Carrel (celui-ci mourut en 1944 avant de pouvoir partager leur succès). Ils disposaient cette fois d'une arme nouvelle et puissante contre les bactéries contaminant les cultures de tissus : les antibiotiques. En ajoutant de la pénicilline et de la streptomycine à la réserve de sang qui maintenait les tissus en vie, ils s'aperçurent qu'on pouvait y faire pousser les virus sans problème. Ils essayèrent aussitôt de travailler avec le virus de la poliomyélite et à leur plus grande joie, celui-ci se développait très bien dans ce milieu. Ce fut la solution miracle qui devait permettre de vaincre la polyomyélite, et les trois hommes reçurent le prix Nobel de médecine et physiologie en 1954.

Le virus de la poliomyélite pouvait maintenant être cultivé en tube à essai, ce qui évitait d'avoir à utiliser des singes (un matériel de laboratoire coûteux et fort capricieux), et l'expérimentation à grande échelle avec le virus devenait possible. Grâce aux techniques de cultures de tissus, Jonas Edward Salk, de l'université de Pittsburgh, put expérimenter divers traitements chimiques sur le virus, et découvrir que le virus de la poliomyélite tué par le formaldéhyde restait capable de produire des réactions immunes dans le corps ; il inventa ainsi le *vaccin de Salk*, aujourd'hui bien connu.

La grande mortalité due à la poliomyélite, la paralysie redoutable qu'elle provoque, son aptitude à frapper des enfants (qui fait qu'on l'appelle aussi *paralysie infantile*), son apparition récente dans le monde moderne puisqu'on ne signale pas d'épidémie avant 1840, et enfin l'intérêt apporté à cette maladie à cause d'une victime éminente, Franklin Delano Roosevelt, tout cela fit de sa conquête l'une des victoires sur la maladie les plus fêtées de toute l'histoire humaine. Aucune nouvelle médicale ne reçut jamais une ovation comparable à celle qui eut lieu en 1955 quand le comité d'évaluation chargé de l'enquête déclara que le vaccin Salk était efficace ; ce fut comme une première à Hollywood. Il est vrai que cette victoire méritait bien une telle publicité. Mais la science ne tire pas avantage des excès de popularité et de publicité. La pression exercée par le public, désireux d'utiliser le vaccin, aboutit à ce que quelques échantillons du vaccin probablement défectueux et dangereux se glissèrent dans le lot, et finalement à une pression du public en sens inverse qui fit reculer d'autant le programme

de vaccination contre la maladie. Ce recul fut finalement comblé, et le vaccin Salk fut à nouveau considéré efficace et sans danger lorsqu'il était correctement préparé. En 1957, le microbiologiste américain d'origine polonaise Albert Bruce Sabin fit un pas de plus en se servant non pas du virus mort (qui peut être dangereux quand il n'est pas tout à fait mort), mais de souches vivantes du virus, incapables de produire la maladie elle-même, mais capables de déterminer la production des anticorps appropriés. Le vaccin, appelé *vaccin de Sabin,* peut de plus être ingéré oralement et ne nécessite pas d'injection hypodermique. Le vaccin de Sabin, d'abord en usage en Union Soviétique puis dans les pays de l'Europe de l'Est, devint d'usage courant aux États-Unis vers 1960, où la crainte de la poliomyélite a disparu depuis.

LES ANTICORPS

Quel est l'effet exact d'un vaccin ? Dans la réponse à cette question réside probablement la clé du problème que l'immunité pose en termes de chimie.

Depuis maintenant plus d'un demi-siècle, les biologistes savent que les principaux moyens de défense du corps contre les infections sont les *anticorps.* Nous laissons de côté pour le moment les globules blancs, appelés *phagocytes,* qui ont la propriété d'ingérer les bactéries – comme l'a démontré en 1883 le biologiste russe Ilya Ilitch Mechnikov, successeur de Pasteur à la tête de l'Institut Pasteur à Paris, et qui partagea en 1908 le prix Nobel de médecine et de physiologie avec Ehrlich. Ces phagocytes ne servent pas à lutter contre les virus et ne sont probablement pas impliqués dans le processus d'immunité décrit ci-dessous. Rappelons qu'on appelle *antigène* toute substance chimique étrangère (par exemple un virus) qui pénètre dans le corps, et anticorps la substance fabriquée par le corps pour combattre cet antigène spécifique qu'il inhibe en se combinant à lui.

Bien avant que les chimistes aient achevé d'analyser les anticorps, ils savaient déjà que c'étaient des protéines ; en effet, des antigènes de nature protéique étaient déjà connus, et seule une protéine pouvait avoir une structure propre à s'associer à une autre protéine, distinguer un antigène spécifique et se combiner avec lui.

Landsteiner (l'inventeur des groupes sanguins) réalisa au début des années vingt une série d'expériences démontrant clairement la spécificité des anticorps. Les substances qu'il utilisait pour entraîner la formation d'anticorps n'étaient pas des antigènes mais des composés beaucoup plus simples de structure chimique connue, à savoir des composés à base d'arsenic appelés *acides arsaniliques.* Ces acides arsaniliques agissent comme des antigènes et entraînent la formation d'anticorps dans le sérum sanguin lorsqu'ils sont injectés dans un animal. Ces anticorps sont spécifiques de l'acide arsanilique utilisé, comme le montre le fait que le sérum sanguin de l'animal ne réagit qu'avec la combinaison acide arsanilique – albumine et non avec l'albumine seule. En fait, on peut même parfois faire réagir les anticorps avec l'acide arsanilique tout seul, sans qu'il soit combiné avec de l'albumine. Landsteiner a aussi montré que de toutes petites modifications dans la structure de l'acide arsanilique sont reflétées dans la structure de l'anticorps, et que l'anticorps provoqué par un type d'acide arsanilique ne réagit pas avec une variété légèrement différente.

Landsteiner inventa le mot *haptène* (du mot grec signifiant « se lier à ») pour désigner les composés, tels que les acides arsaniliques, qui engendrent la fabrication d'anticorps lorsqu'ils sont combinés avec une protéine. Il y a probablement pour chaque antigène naturel une région spécifique dans la molécule qui agit comme un haptène. Selon cette théorie, pour qu'une bactérie ou un virus devienne utilisable comme vaccin, il faut que sa structure soit suffisamment modifiée pour diminuer sa capacité à endommager les cellules, mais que son groupe haptène reste intact de façon à provoquer la formation d'un anticorps spécifique. C'est le cas du vaccin contre la grippe, que le groupe de Richard A. Lerner a mis au point au début des années quatre-vingts à l'aide d'une protéine synthétique fabriquée d'après celle du virus de la grippe, et qui est efficace en laboratoire sur des cobayes.

L'intérêt de connaître la nature chimique des haptènes naturels est que, si celle-ci peut être déterminée, on peut en faire usage en combinaison avec une protéine inoffensive comme vaccin pour provoquer des anticorps contre un antigène spécifique. On évite ainsi d'avoir recours à des toxines ou à des virus atténués, méthodes qui comportent toujours un certain risque.

Quel est le mécanisme exact de la production d'un anticorps suscité par un antigène ? Peut-être, comme le pensait Ehrlich, le corps contient-il à l'état normal une petite quantité de tous les anticorps dont il a besoin et l'antigène étranger, en réagissant avec l'anticorps approprié, stimule le corps à produire des quantités supplémentaires de cet anticorps. Certains immunologistes sont encore partisans de cette théorie légèrement modifiée. Toutefois il semble très improbable que le corps soit doté d'anticorps spécifiques pour tous les antigènes possibles, y compris des substances aussi peu naturelles que les acides arsaniliques.

Selon une autre théorie, le corps contient une sorte de molécule très plastique dont la conformation s'adapte d'abord à n'importe quel antigène ; l'antigène sert alors de matrice pour déterminer la forme d'un anticorps spécifique fabriqué en réponse à celui-ci. Pauling a proposé ainsi en 1940 que les différents anticorps seraient des versions différentes de la même molécule de base, repliée sur elle-même de différentes façons ; en d'autres termes, la conformation de l'anticorps s'adapterait à son antigène, de même qu'un gant s'adapte à une main.

Quoi qu'il en soit, en 1969, les progrès en analyse de protéines ont rendu possible à l'équipe de Gerald Maurice Edelman la détermination de la séquence en acides aminés d'un anticorps type, séquence constituée de plus de mille acides aminés. Edelman a reçu pour son travail le prix Nobel de médecine et physiologie en 1972.

Ce travail a été repris par J. Donald Capra, qui a montré qu'il existait des *régions hypervariables* dans la séquence en acides aminés des anticorps. Il semble que les portions les plus constantes de la chaîne servent à former une structure tridimensionnelle qui maintient la région hypervariable, laquelle est elle-même conçue pour s'adapter à un antigène spécifique, à la fois par des changements dans des acides aminés donnés à l'intérieur de la chaîne et par des changements dans la configuration géométrique de la molécule.

C'est à l'aide de cette aptitude à se combiner avec l'antigène qu'un anticorps peut neutraliser une toxine et rendre impossible sa participation à des réactions nuisibles dans le corps, ou se combiner avec des régions

de la surface d'un virus ou d'une bactérie. Quand l'anticorps a la capacité de se combiner avec deux endroits différents, dont l'un à la surface d'un micro-organisme et l'autre à la surface d'un autre, il fait démarrer un processus d'agglutination par lequel les micro-organismes se collent entre eux et perdent leur capacité à se multiplier ou à entrer dans une cellule.

De plus, l'association avec des anticorps sert aussi à identifier les cellules portant l'antigène de sorte que les phagocytes les ingèrent plus facilement, et à activer le système du complément qui, en faisant usage de ses composants enzymatiques, perce la paroi de la cellule importune et la détruit.

A la limite, la spécificité des anticorps peut être un désavantage d'un certain point de vue. Supposez par exemple qu'un virus soit muté de telle sorte que sa protéine ait une structure légèrement différente : l'ancien anticorps contre ce virus ne peut plus s'adapter à la nouvelle structure, et par conséquent l'immunité contre une souche virale ne protège pas nécessairement contre une autre souche. Les virus de la grippe et du rhume sont particulièrement sujets à des mutations mineures – d'où la réapparition fréquente de ces maladies. La grippe en particulier trouve de temps en temps le moyen de développer des mutants très virulents qui font alors des catastrophes dans une population non immunisée (ce fut le cas en 1918, et à un moindre degré lors de l'épidémie pandémique de *grippe asiatique* en 1957).

Un autre désavantage de la trop grande efficacité du corps à développer des défenses immunes est son aptitude à produire des anticorps même contre une protéine inoffensive introduite par hasard dans le corps ; celui-ci devient alors *sensibilisé* à cette protéine et peut réagir violemment à toute incursion ultérieure de la protéine originellement inoffensive. La réaction prend alors la forme de démangeaisons, pleurs, production de mucus dans le nez et la gorge, asthme et ainsi de suite. Cette sorte de *réaction allergique* est provoquée par le pollen de certaines plantes (causant le rhume des foins), certains aliments, la fourrure d'animaux, etc. La réaction allergique peut être grave au point de provoquer une invalidité et même la mort. La découverte des *chocs anaphylactiques* par le Français Charles Robert Richet lui valut le prix Nobel de médecine et de physiologie en 1913.

En un sens, on peut dire que chaque être humain est allergique aux cellules de tous les autres êtres humains. En effet, une transplantation ou une greffe d'un individu à un autre ne prend pas parce que le corps du receveur traite le tissu transplanté comme une protéine étrangère et fabrique des anticorps contre celui-ci. Les greffes les plus réussies sont celles qui ont lieu entre deux jumeaux, parce que leur hérédité commune leur donne exactement les mêmes protéines, et ils peuvent ainsi échanger des tissus, voire un organe entier (par exemple le rein).

La première transplantation rénale réussie a eu lieu en décembre 1954 à Boston, entre deux jumeaux. Le receveur est mort en 1962 à l'âge de trente ans d'une maladie des artères coronaires. Depuis, des centaines d'individus ont survécu pendant des mois et même des années avec des reins transplantés. Des essais destinés à transplanter d'autres organes, tels que les poumons ou le foie, ont été faits, mais l'opinion publique a surtout été marquée par la transplantation cardiaque. Les premières transplantations cardiaques réussies ont été réalisées en décembre 1967 par le chirurgien sud-africain Christian Barnard. Le receveur – Philip Blaiberg, un dentiste sud-africain à la retraite – a vécu pendant plusieurs mois avec le cœur d'un autre être humain.

Les transplantations cardiaques ont fait fureur pendant quelque temps, mais leur nombre avait bien diminué fin 1969 ; en effet, peu de receveurs vivent longtemps et les problèmes de rejets de tissus sont apparemment insolubles, en dépit des tentatives renouvelées pour vaincre la répugnance du corps à incorporer des tissus autres que les siens.

C'est le bactériologiste australien Macfarlane Burnet qui a suggéré le premier que les tissus embryonnaires peuvent être *rendus immunes* contre les tissus étrangers, et que l'animal tolère mieux ensuite les greffes de tissus. Cette hypothèse a été démontrée par le biologiste anglais Peter Medawar sur des embryons de souris, et les deux hommes ont partagé en 1960 le prix Nobel de médecine et de physiologie.

En 1962 l'immunologiste franco-australien Jacques Francis Albert Pierre Miller, qui travaillait en Angleterre, alla plus loin et découvrit pourquoi il est possible de travailler avec des embryons pour obtenir la tolérance des greffes. Il montra que la glande du thymus (un organe dont l'usage, jusque-là, était inconnu) était le tissu capable de former des anticorps. Lorsqu'on enlève cette glande à la naissance à une souris, celle-ci meurt au bout de trois ou quatre mois par suite de son incapacité à se protéger contre l'environnement. Par contre, quand on laisse subsister le thymus pendant trois semaines dans la souris, l'organe a le temps de provoquer le développement de cellules productrices d'anticorps et peut ensuite être enlevé sans danger pour la souris. De même, on peut traiter les embryons dans lesquels le thymus n'a pas encore fait son effet pour leur apprendre à tolérer les tissus étrangers ; on peut espérer à l'heure actuelle améliorer la tolérance des tissus aux greffes grâce au rôle du thymus, et ceci même pour des tissus adultes, si nécessaire.

Il subsiste néanmoins des problèmes à résoudre, une fois celui du rejet de tissus surmonté. Par exemple, pour chaque personne qui reçoit un organe vivant, il faut trouver une autre personne pour le fournir, d'où la question du moment où le donneur potentiel peut être considéré comme mort et comme pouvant céder l'un de ses organes.

A cause de cela, on peut préférer préparer des organes artificiels qui n'impliquent pas de rejets de tissus, ni de problèmes éthiques complexes. Les reins artificiels, par exemple, sont devenus d'usage courant dès 1940, et les malades qui n'ont pas de fonctions rénales naturelles peuvent se rendre dans des hôpitaux une ou deux fois par semaine pour que leur sang soit nettoyé de ses déchets. C'est là une vie très limitée, même lorsqu'on a la chance d'être soigné, mais tout de même préférable à la mort.

C'est vers 1940 qu'on a découvert une cause aux réactions allergiques : elles seraient provoquées par la libération dans le flux sanguin de petites quantités d'une substance appelée *histamine*. On a pu alors produire des *antihistamines* neutralisantes, capables de soulager les symptômes allergiques sans bien sûr supprimer l'allergie elle-même. La première antihistamine réussie a été produite à l'Institut Pasteur de Paris en 1937 par le chimiste suisse Daniel Bovet, qui a reçu en 1957 le prix Nobel de médecine et de physiologie pour ce résultat et d'autres recherches en chimiothérapie.

Les firmes pharmaceutiques, constatant que les reniflements et autres symtômes allergiques ressemblaient à ceux des rhumes, ont tenté d'élargir le champ d'application des antihistamines et inondé le pays de ces comprimés entre 1949 et 1950. (La mode de ces comprimés, en fait très peu efficaces contre les rhumes, a diminué par la suite.)

Les allergies sont surtout nocives quand le corps devient allergique à l'une ou l'autre de ses propres protéines. D'habitude, le corps s'ajuste à ses protéines au cours de son développement à partir d'un œuf fertilisé ; mais il arrive qu'une partie de cet ajustement se perde ; soit parce que le corps s'immunise contre une protéine étrangère, trop proche par sa structure de celle d'une de ses protéines ; soit parce qu'avec l'âge, des changements ont lieu à la surface de certaines cellules, qui apparaissent comme étrangères aux cellules productrices d'anticorps ; soit encore parce que des virus inconnus, qui ne font que peu ou pas de tort à une cellule qu'ils infectent, produisent malgré tout des changements subtils à sa surface. Le résultat en est les *maladies auto-immunes.*

En tout cas les réponses auto-immunes interviennent plus fréquemment dans les maladies humaines qu'on ne le pensait jusqu'à présent. La plupart des maladies auto-immunes sont rares, en dehors de l'arthrite rhumatoïde qui est peut-être une maladie auto-immune. Le traitement de ce genre de maladies est difficile, mais bien sûr on a de meilleures chances de trouver la direction dans laquelle chercher un traitement efficace contre une maladie si on en connaît la cause.

En 1937, grâce aux techniques d'isolement des protéines par électrophorèse, les biologistes ont finalement réussi à localiser les anticorps dans le sang, dans une fraction sanguine appelée *gammaglobulines.*

Les médecins savent depuis longtemps que certains enfants ne peuvent pas fabriquer des anticorps et sont par conséquent une proie facile pour l'infection. En 1951, les médecins de l'hôpital Walter Reed à Washington, au cours de l'analyse électrophorétique du plasma d'un garçon de huit ans souffrant de *septicémie* grave (« empoisonnement de sang »), remarquèrent à leur grand étonnement que son sang ne contenait pas du tout de gammaglobulines. D'autres cas furent bientôt découverts. Les chercheurs établirent que ce défaut métabolique était inné et privait la personne de l'aptitude à fabriquer des gammaglobulines ; cette maladie s'appelle *agammaglobulinémie.* Les sujets concernés ne peuvent pas développer d'immunité contre les bactéries. On peut toutefois les maintenir en vie, à l'heure actuelle, grâce à des antibiotiques. Si surprenant que ce soit, ils sont capables de devenir immunes aux infections virales, comme la rougeole ou la varicelle, s'ils ont attrapé la maladie une fois, ce qui semble prouver que les anticorps ne sont pas le seul moyen de défense possible du corps contre les virus.

En 1957, un groupe de bactériologistes anglais, dirigé par Alick Isaacs, a montré que sous l'effet d'une invasion virale, les cellules infestées produisent une protéine ayant des propriétés antivirales à large spectre d'action. Celle-ci neutralise non seulement le virus impliqué dans l'infection immédiate, mais également un second virus non apparenté au premier qui infecte le même animal. Cette protéine, appelée *interféron,* est produite plus rapidement que les anticorps ne le sont et peut expliquer les défenses antivirales des gens qui souffrent d'agammaglobulinémie. Il semble que sa production soit stimulée par la présence d'ARN double brin du type que l'on trouve dans les virus animaux ; peut-être l'interféron dirigerait-il la synthèse de messager produisant une protéine antivirale, qui elle-même inhiberait la production des protéines virales. L'interféron est aussi puissant que les antibiotiques et n'active pas de résistance. Il est par contre très spécifique de l'espèce : seul l'interféron produit par des êtres humains et par d'autres primates peut agir sur les cellules des êtres humains.

Le fait que de l'interféron humain soit nécessaire et que les cellules humaines ne le produisent qu'à l'état de traces a rendu impossible pendant longtemps l'obtention d'un matériel en quantité suffisante pour l'usage clinique.

Sydney Pestka, à l'institut Roche de biologie moléculaire, s'est mis à travailler sur des méthodes de purification de l'interféron en 1977. Il a découvert que l'interféron existait sous forme de plusieurs protéines de structures proches. L'interféron-alpha, le premier à avoir été purifié, a un poids moléculaire de 17 500 daltons, et il est composé d'une chaîne de 170 acides aminés. La séquence en acides aminés de l'interféron d'une douzaine d'espèces différentes est connue, et ces séquences ne présentent que des différences mineures entre elles.

Le gène responsable de la formation de l'interféron est maintenant connu, et il a été inséré dans la bactérie *Escherichia coli* par le biais des méthodes de génie génétique. Une colonie de ces bactéries a été induite à fabriquer de l'interféron humain sous une forme très pure, qu'on a pu isoler et cristalliser. Les cristaux ont été analysés aux rayons X, et la structure tridimensionnelle déterminée.

En 1981, on disposait de suffisamment d'interféron pour réaliser des essais cliniques. Aucun miracle ne s'est produit, mais avec le temps un procédé approprié sera sûrement mis au point.

De nouvelles maladies infectieuses font leur apparition de temps en temps. Les années quatre-vingts ont vu apparaître une maladie effrayante appelée le SIDA *(syndrôme d'immuno-déficience acquise)*, dans laquelle les mécanismes d'immunisation sont détruits et où une seule infection suffit à tuer. Cette maladie, qui affecte surtout des homosexuels masculins, des Haïtiens et des polytransfusés, se répand rapidement et elle est généralement mortelle. Elle reste pour l'instant incurable, mais le virus qui en est la cause a été isolé en France puis aux États-Unis en 1984, ce qui représente un progrès important.

Le cancer

Si le danger des maladies infectieuses a diminué, la fréquence des autres types de maladies a en revanche augmenté. Bien des gens qui, il y a un siècle, seraient morts jeunes de la tuberculose, de la diphtérie, de la pneumonie ou du typhus, vivent maintenant assez longtemps pour mourir d'une maladie cardiaque ou du cancer. C'est l'une des raisons pour lesquelles ces deux maladies sont devenues respectivement les numéros un et deux des causes de décès dans le monde occidental. Le cancer est devenu une véritable menace planant au-dessus de nos têtes, prête à frapper n'importe qui, et il a pour ainsi dire pris la place de la peste et de la variole dans l'imagination collective. Trois cent mille Américains en meurent chaque année, et l'on recense dix mille nouveaux cas chaque semaine, soit une augmentation de 50 % par rapport à 1900.

Le cancer est en fait tout un ensemble de maladies (environ deux cents types connus), qui affectent les différentes parties du corps de diverses façons. Mais la manifestation principale est toujours la même :

désorganisation et croissance incontrôlée des tissus affectés. Le nom *cancer* (du mot latin pour « crabe ») vient du fait qu'Hippocrate et Galien imaginaient que cette maladie se propageait à travers les veines malades comme les pinces crochues et déployées d'un crabe.

Le mot *tumeur* (du mot latin qui signifie « pousser ») n'est pas synonyme de cancer car il s'applique à des excroissances sans danger comme les verrues et les grains de beauté (*tumeurs bénignes*) aussi bien qu'aux véritables cancers (*tumeurs malignes*). Les cancers portent des noms différents selon les tissus affectés. Les cancers de la peau ou des villosités intestinales (les maladies malignes les plus fréquentes) sont appelés *carcinomes* (du mot grec pour « crabe ») ; les cancers des tissus conjonctifs sont des *sarcomes,* ceux du foie des *hépatomes,* ceux des glandes en général des *adénomes,* ceux des globules blancs des *leucémies ;* et ainsi de suite.

L'Allemand Rudolf Virchow, le premier à avoir étudié un tissu cancéreux sous le microscope, pensait que le cancer était provoqué par des irritations ou des chocs dus à l'environnement extérieur. Cette idée était toute naturelle, car ce sont justement les parties du corps les plus exposées au monde extérieur qui sont les plus sujettes au cancer. Mais avec la théorie bactérienne des maladies, les pathologistes ont commencé à chercher un microbe capable de provoquer le cancer. Virchow était un opposant résolu de la théorie bactérienne des maladies, et insistait vigoureusement sur la théorie de l'irritation. A vrai dire, il quitta la pathologie pour faire de l'archéologie et de la politique, quand il s'avéra que la théorie bactérienne des maladies allait gagner.

Si les raisons pour lesquelles Virchow s'entêtait n'étaient pas exactes, il avait peut-être raison d'un certain point de vue ; les preuves s'accumulaient que certains environnements étaient particulièrement favorables au cancer. Au XVIIIe siècle, on observa que les ramoneurs étaient plus souvent sujets au cancer du scrotum que d'autres personnes, et après que les colorants à base de goudron furent développés, on découvrit que chez les ouvriers des industries de colorants, le taux de cancers de la peau ou de la vessie était supérieur à la moyenne. Il semble qu'un élément contenu dans la suie et dans les colorants aniliques soit capable de causer le cancer. Puis en 1915, deux scientifiques japonais, K. Yamagiwa et K. Ichikawa, remarquèrent qu'un composé du goudron pouvait produire des cancers chez le lapin lorsqu'on l'appliquait à ses oreilles pendant un certain temps. En 1930, deux chimistes anglais induisirent des cancers dans des animaux avec un produit chimique synthétique appelé *dibenzanthracène* (un hydrocarbure dont la molécule est composée de cinq anneaux de benzène). Ce composé ne figure pas dans le goudron, mais trois ans plus tard on découvrit que le *benzopyrène* (qui contient aussi cinq anneaux benzéniques, mais d'un arrangement différent), un produit chimique qui apparaît dans le goudron, provoquait effectivement le cancer.

Un grand nombre de *carcinogènes* (agents produisant des cancers) ont maintenant été identifiés. Beaucoup d'entre eux sont des hydrocarbures composés de nombreux anneaux benzéniques, comme les deux premiers qui furent découverts. Certains sont des molécules apparentées aux colorants aniliques. En fait, l'un des soucis principaux concernant l'usage des colorants artificiels dans la nourriture est la crainte que, consommés abusivement, ces colorants puissent devenir des carcinogènes.

Beaucoup de biologistes pensent que les êtres humains ont introduit au cours des deux ou trois derniers siècles un certain nombre de facteurs

nouveaux qui engendrent le cancer dans l'environnement. Il y a l'usage intensif du charbon, la combustion du pétrole à grande échelle, en particulier dans les moteurs à essence, l'utilisation croissante des produits chimiques synthétiques dans la nourriture, les cosmétiques et ainsi de suite. Le suspect le mieux identifié bien sûr est la fumée des cigarettes, qui s'accompagne d'un taux relativement élevé du cancer des poumons.

L'EFFET DES RADIATIONS

Parmi les facteurs de l'environnement dont on est sûr qu'ils sont carcinogènes, se trouvent les radiations énergétiques, auxquelles les êtres humains ont été exposés progressivement depuis 1895.

Le 5 novembre 1895, le physicien allemand Wilhelm Konrad Roentgen réalisa une expérience pour étudier la luminescence produite par les rayons cathodiques. Il éteignit la pièce pour mieux en voir l'effet ; quand il alluma le tube à rayons cathodiques, qui était enfermé dans une boîte de carton noire, il eut la surprise d'apercevoir un éclair de lumière provenant de l'autre côté de la pièce. L'éclair provenait d'une feuille de papier enduite de platinocyanure de baryum, un produit chimique luminescent. Etait-il possible que les radiations de la boîte fermée aient induit ce rayonnement ? Roentgen éteignit son tube à rayons cathodiques et la lueur s'arrêta. Il le ralluma et la lueur réapparut. Il emporta le papier dans une autre pièce où le rayonnement continuait toujours. De toute évidence, le tube à rayons cathodiques produisait une forme de radiations pouvant traverser le carton et les murs.

Roentgen, n'ayant aucune idée de la sorte de radiations que cela pouvait être, les appela simplement rayons X. D'autres scientifiques essayèrent de changer le nom en *rayons Roentgen,* mais ce mot était si difficile à prononcer pour ceux qui n'étaient pas de langue allemande que c'est le nom de *rayons X* qui subsista. Nous savons maintenant que les électrons accélérés qui constituent les rayons cathodiques sont fortement décélérés quand ils touchent une barrière métallique. L'énergie cinétique perdue est convertie en radiations appelées *Bremsstrahlung* – allemand pour « radiations de freinage ». Les rayons X sont un exemple de ce genre de radiations.

Les rayons X ont révolutionné la physique, ils ont envahi l'imagination des physiciens, fait naître des quantités d'expériences nouvelles, amené en quelques mois à la découverte de la radioactivité et ouvert le monde intérieur de l'atome. Quand les récompenses des prix Nobel commencèrent en 1901, Roentgen fut le premier à recevoir le prix en physique.

Mais les rayons X durs ont fait débuter autre chose – à savoir l'exposition des êtres humains à des intensités de radiations hautement énergétiques telles qu'ils ne l'avaient jamais vécu avant. Quatre jours après que la nouvelle de la découverte de Roentgen fut parvenue aux États-Unis, les rayons X étaient utilisés pour localiser une balle dans la jambe d'un patient. C'était en effet un moyen merveilleux d'explorer l'intérieur du corps ; les rayons X passent aisément à travers les tissus mous (constitués principalement d'éléments de faible poids moléculaire) et ont tendance à être arrêtés par les éléments de poids moléculaire plus élevé comme ceux qui constituent les os (composés en grande partie de phosphore et de calcium). Sur une plaque photographique placée derrière le corps, on voit les os d'un blanc opaque, en contraste avec les surfaces noires que les rayons X ont

traversées en plus grande intensité parce qu'ils ont été beaucoup moins absorbés par les tissus mous. Une balle de plomb est complètement blanche car elle arrête totalement les rayons X.

Les rayons X sont d'une utilité évidente pour détecter les fractures osseuses, les articulations calcifiées, les trous dans les dents, les objets étrangers dans le corps, et ainsi de suite. Il est également facile de faire ressortir les tissus mous en introduisant dedans le sel insoluble d'un élément lourd. Le sulfate de baryum, lorsqu'il est avalé, dessine l'estomac ou les intestins. Un composé iodé injecté dans le sang se déplace jusqu'aux reins et l'urètre et permet de voir ces organes, parce que l'iode a un poids moléculaire élevé, et par conséquent il est opaque aux rayons X.

Même avant que les rayons X aient été découverts, un médecin danois, Niels Ryberg Finsen, avait remarqué que les radiations à haute énergie pouvaient tuer les micro-organismes ; il se servait de lumière UV pour détruire une bactérie causant une maladie de la peau appelée *lupus vulgaris* (il reçut pour cela le prix Nobel de physiologie et de médecine en 1903). Les rayons X se révélèrent bien plus létaux, pouvant tuer le micro-organisme qui provoque la teigne tonsurante ; ils peuvent endommager ou détruire les cellules humaines et sont en fin de compte utilisés pour anéantir les cellules cancéreuses que le bistouri du chirurgien ne peut atteindre.

Ce qui a été aussi découvert – à grand-peine – c'est que les radiations à haute énergie peuvent *provoquer* des cancers. Au moins une centaine des premières personnes à avoir travaillé avec des rayons X et des matériaux radioactifs sont mortes du cancer, la première mort datant de 1902. Ainsi, aussi bien Marie Curie que sa fille Irène Joliot-Curie moururent de leucémie, et il est facile de penser que les radiations y ont contribué dans les deux cas. En 1928, un médecin anglais, George William Marshall Findlay, a observé que même les radiations ultraviolettes sont suffisamment énergétiques pour provoquer des cancers de la peau chez la souris.

On a probablement raison de soupçonner que l'exposition croissante des êtres humains à des radiations énergétiques (sous la forme de rayons X médicaux, d'expérimentations nucléaires et ainsi de suite) est responsable en partie de l'augmentation croissante du cancer.

MUTAGÈNES ET ONCOGÈNES

Que peuvent bien avoir en commun les divers carcinogènes – produits chimiques, radiations, et autres ? Il semble qu'ils puissent tous provoquer des mutations génétiques, et que le cancer soit le résultat de mutations dans les cellules du corps. C'est le zoologue allemand Theodor Boveri qui a suggéré en premier cette idée en 1914.

Supposons par exemple qu'un gène donné soit modifié de sorte qu'il ne puisse plus produire une enzyme nécessaire au processus de contrôle de la croissance des cellules. Quand une cellule possédant un tel gène défectueux se divise, elle transmet le défaut à ses descendants. A cause de ce défaut dans le mécanisme de contrôle qui ne fonctionne pas, les divisions supplémentaires de ces cellules peuvent continuer indéfiniment, sans respecter les besoins du corps en tant que tel ou même ceux des tissus impliqués (par exemple, la spécialisation des cellules en un organe). Le tissu est désorganisé, et pour ainsi dire dans un état d'anarchie par rapport au corps.

Que les radiations à haute énergie puissent produire des mutations est bien établi. Qu'en est-il des carcinogènes chimiques ? Là aussi, les mutations par les produits chimiques ont été démontrées. Les *moutardes azotées* sont un exemple très clair. Ces composés, comme le *gaz moutarde* utilisé pendant la Première Guerre mondiale, produisent des brûlures et des cloques sur la peau qui ressemblent à celles que provoquent les rayons X ; ils peuvent aussi endommager les chromosomes et augmenter le taux de mutations. De plus, bien d'autres produits chimiques ont été découverts qui imitent les radiations énergétiques de ce point de vue.

Les produits chimiques qui induisent des mutations sont appelés des *mutagènes.* Tous les mutagènes ne sont pas des carcinogènes, et tous les carcinogènes ne sont pas des mutagènes. Mais il existe suffisamment d'exemples de composés à la fois carcinogènes et mutagènes pour faire naître l'hypothèse d'une relation de cause à effet.

A partir des années soixante, les scientifiques ont commencé à chercher si des changements qui n'étaient pas dus au hasard avaient lieu dans les chromosomes des cellules tumorales, par comparaison avec les cellules normales. En effet, ils en ont trouvé qui ont été mis en évidence plus précisément grâce aux techniques développées pour former des cellules hybrides entre espèce humaine et souris. Ces cellules hybrides ne contiennent que relativement peu de chromosomes humains ; si l'un de ceux que l'on suspecte d'intervenir dans l'activation des tumeurs est inclus, il donne naissance à une tumeur quand la cellule hybride est injectée dans la souris.

De plus amples investigations ont associé le changement cancéreux à un seul gène sur un chromosome, quand un groupe du Massachusetts Institute of Technology, dirigé par Robert A. Weinberg, produisit avec succès des tumeurs dans la souris en 1978 par le transfert de gènes individuels. Ceux-ci ont été appelés *oncogènes* (le préfixe *onco-,* du mot grec signifiant « amas », est d'usage courant en terminologie médicale pour « tumeur »).

Un oncogène est tout à fait semblable à un gène normal. Les deux diffèrent probablement par un seul acide aminé le long de la chaîne. On pense donc à un *proto-oncogène,* un gène normal présent dans les cellules et transféré à chaque division cellulaire qui, sous l'effet d'influences diverses, peut, à un moment donné, subir un changement qui en fait un oncogène actif. Bien sûr, on peut se demander pourquoi il y a un proto-oncogène libre dans la cellule quand le danger potentiel est si grand. Pour l'instant nous n'avons pas de réponse, mais nous avons au moins une nouvelle direction de recherche, ce qui est appréciable.

LA THÉORIE DES VIRUS

L'idée selon laquelle des micro-organismes provoqueraient le cancer est loin d'avoir disparu pendant ce temps. Avec la découverte des virus, cette suggestion de l'ère pasteurienne a repris. En 1903, le bactériologiste français Amédée Borrel a émis l'hypothèse que le cancer était peut-être une maladie virale ; et en 1908 deux Danois, Wilhelm Ellerman et Olaf Bang, ont montré que la leucémie des oiseaux était effectivement causée par un virus. Toutefois, la leucémie n'était à ce moment-là pas reconnue comme une forme de cancer, et la question ne fut pas débattue plus avant. En 1909,

le médecin américain Francis Peyton Rous décida de broyer une tumeur de poulet, de la filtrer et d'injecter le filtrat clair dans d'autres poulets. Certains d'entre eux développèrent des tumeurs. Plus le filtre était fin, moins il y avait de tumeurs. Il paraissait évident que des particules d'une sorte ou d'une autre étaient responsables des tumeurs, et ces particules semblaient être de la taille des virus.

Les *virus tumorigènes* ont une histoire complexe. Tout d'abord les tumeurs pouvant être associées à des virus se révélèrent uniformément bénignes ; par exemple, on montra que des virus provoquent les papillomes du lapin (semblables aux verrues). Mais en 1936, John Joseph Bittner, qui travaillait dans le célèbre laboratoire d'élevage de souris à Bar Harbor, Maine, tomba sur quelque chose de plus important. Maude Slye, du même laboratoire, avait élevé des souches de souris présentant apparemment une résistance innée au cancer, et d'autres souches qui étaient plus enclines à la maladie. Les souris de certaines souches développaient rarement des cancers, celles des autres souches en développaient toujours à maturité. Bittner tenta d'échanger les mères des souriceaux nouveau-nés de sorte que ceux-ci têtent le lait d'une souche opposée à la leur. Il découvrit que lorsqu'un souriceau d'une souche résistante au cancer était nourri par des mères d'une souche non résistante au cancer, il développait généralement un cancer. D'un autre côté, les souriceaux enclins au cancer nourris par des mères résistantes au cancer ne le développaient pas. Bittner en conclut que la cause du cancer, quelle qu'elle soit, n'était pas innée mais transmise par le lait de la mère. Il appela cet agent *facteur du lait.*

Bien entendu, on suspectait le facteur du lait de Bittner d'être un virus. Finalement, le biochimiste Samuel Graff de l'université Columbia identifia le facteur comme une particule contenant des acides nucléiques. D'autres virus tumoraux, capables de provoquer certains types de tumeurs chez la souris ainsi que des leucémies animales, ont été trouvés et presque tous contiennent des acides nucléiques. Aucun virus n'a été détecté en connection avec des cancers humains, mais la recherche sur les cancers humains a des limites évidentes.

A présent, les théories mutationnelles et virales du cancer commencent à converger. Peut-être les contradictions apparentes entre ces deux notions n'en sont-elles pas. Les virus et les gènes se ressemblent beaucoup par leur structure, et certains virus, lorsqu'ils envahissent la cellule, peuvent s'intégrer à l'équipement cellulaire et jouer le rôle d'oncogènes.

Bien sûr, les virus tumoraux sont toujours des virus à ARN, tandis que les gènes humains possèdent de l'ADN. Tant que l'on considérait comme évident que l'information circulait toujours de l'ADN vers l'ARN, il était difficile de voir comment les virus tumoraux pouvaient jouer le rôle des gènes. On sait maintenant que, dans certains cas, l'ARN peut provoquer la production d'ADN, lequel véhicule le schéma des nucléotides de l'ARN. Un virus tumoral n'est donc peut-être pas lui-même un oncogène, mais il peut *fabriquer* un oncogène.

D'ailleurs, l'attaque virale peut être moins directe. Le virus peut simplement jouer un rôle important dans la conversion d'un proto-oncogène en oncogène.

Ce n'est qu'en 1966 que l'hypothèse des virus a été estimée suffisamment féconde pour mériter un prix Nobel. Peyton Rous, qui avait fait sa découverte cinquante-cinq ans plus tôt, était heureusement encore en vie et reçut une part du prix Nobel de médecine et de physiologie en 1966

(il vécut jusqu'en 1970, et mourut à l'âge de 90 ans, faisant encore de la recherche presque jusqu'à la fin).

LES TRAITEMENTS POSSIBLES

Comment la croissance incontrôlée des cellules se manifeste-t-elle au niveau du métabolisme ? Cette question n'a pas encore reçu de réponse. L'une des hypothèses possibles implique des systèmes hormonaux, par exemple les hormones sexuelles.

D'une part, on sait que les hormones sexuelles stimulent la croissance rapide de parties du corps (comme la poitrine chez une jeune adolescente). D'autre part, les tissus des organes sexuels – les seins, le col de l'utérus et les ovaires chez une femme, les testicules et la prostate chez un homme – sont particulièrement vulnérables au cancer. La meilleure preuve de cela est d'ordre chimique. En 1933, le biochimiste allemand Heinrich Wieland (qui avait obtenu le prix Nobel de chimie en 1927 pour son travail sur les acides biliaires) s'arrangea pour convertir un acide biliaire en un hydrocarbure complexe appelé *méthylcholanthrène,* un puissant carcinogène. Or, le méthylcholanthrène (comme les acides biliaires) a la structure à quatre cycles des stéroïdes et de toutes les hormones sexuelles. Est-il possible qu'une molécule d'hormone sexuelle erronée agisse comme un carcinogène ? Ou qu'une hormone correctement conformée soit confondue avec un carcinogène, pour ainsi dire, par un schéma génétique distordu dans la cellule, stimulant ainsi une croissance incontrôlée ? Toutes ces spéculations sont intéressantes mais hypothétiques.

Assez curieusement, un changement dans les réserves d'hormones sexuelles met parfois en échec la croissance cancéreuse. Par exemple la castration, qui réduit la production par le corps d'hormones sexuelles mâles, ou l'administration d'hormones sexuelles femelles de neutralisation, ont un effet atténuant sur le cancer de la prostate. En tant que traitement, on ne peut cependant guère s'en vanter, et c'est une preuve du désespoir provoqué par le cancer qu'on ait dû recourir à de tels procédés.

Le principal moyen de lutter contre le cancer reste la chirurgie. Ses limites sont ce qu'elles ont toujours été : quelquefois le cancer ne peut pas être supprimé sans tuer le patient ; souvent le bistouri libère des morceaux de tissu malin (parce que le tissu cancéreux désorganisé a tendance à se fragmenter), qui sont ensuite transportés dans la circulation sanguine vers d'autres parties du corps où ils se fixent et poussent.

L'utilisation des radiations énergétiques pour tuer le tissu cancéreux a aussi des désavantages. La radioactivité artificielle a ajouté des armes nouvelles aux traditionnels rayons X et au radium. L'une d'entre elles est le cobalt 60, qui émet des rayons gamma à haute énergie et qui est beaucoup moins cher que le radium ; une autre est une solution d'iode radioactif (« cocktail atomique »), qui se concentre sur la glande thyroïde et attaque par conséquent le cancer de la thyroïde. Mais la tolérance du corps aux radiations est limitée et le danger existe toujours que les radiations fassent démarrer plus de cancers qu'elles n'en arrêtent.

Les connaissances croissantes des dernières décennies offrent l'espoir que certaines méthodes de traitement apparaissent qui soient moins énergiques, à la fois plus souples et plus efficaces.

Par exemple, si un virus est impliqué d'une façon ou d'une autre dans l'apparition d'un cancer, tout agent qui inhibe l'action du virus pourrait

diminuer l'incidence du cancer ou arrêter sa croissance une fois qu'il est commencé. On pense immédiatement à l'interféron ; maintenant qu'il peut être obtenu pur et en grande quantité, il a été essayé sur des malades cancéreux ; il n'y a pas eu de succès très net jusqu'à maintenant, mais il s'agit d'un procédé expérimental, testé seulement sur des patients très malades, voire incurables. Peut-être aussi les méthodes utilisées sont-elles à améliorer.

Une autre approche est la suivante : les oncogènes diffèrent peu des gènes normaux ; on peut supposer qu'ils sont formés fréquemment et que la production d'une cellule cancéreuse est plus fréquente que nous ne l'imaginons. Une telle cellule est différente d'une cellule normale par certains aspects et peut-être que le système immune du corps peut la reconnaître tôt et s'en défaire. Dans ce cas, le développement du cancer ne signifie pas qu'une cellule cancéreuse a été formée, mais que celle-ci, une fois formée, n'a pas été arrêtée. Le cancer pourrait alors être le résultat de cette déficience du système immunitaire – d'une certaine façon, l'opposé d'une maladie auto-immune, où le système immune fonctionne trop bien.

La prévention et le traitement du cancer reposent donc sur notre compréhension croissante du fonctionnement de ce système immune. Mais, en attendant qu'un tel but soit atteint, le corps pourrait-il recevoir une aide artificielle sous la forme de composés qui feraient la distinction entre cellules normales et cellules cancéreuses ?

Certaines plantes, par exemple, produisent des substances qui réagissent avec des sucres comme les anticorps réagissent avec les protéines. On ignore pour le moment la raison d'être de ces substances dans les plantes.

Les membranes qui entourent les cellules sont faites de protéines et de corps gras, mais ces protéines incorporent habituellement dans leur structure des molécules de sucres peu complexes. Parce que la nature des sucres de ces membranes varie, les cellules sanguines présentent des typages différents qui se différencient les uns des autres par le fait que certains types s'agglutinent dans certaines conditions et d'autres dans d'autres.

Le biochimiste américain William Clouser Boyd s'est demandé s'il existait des substances végétales qui puissent distinguer les groupes sanguins les uns des autres. En 1954, à sa grande surprise, il trouva une substance de ce type dans les haricots de Lima, une des premières plantes qu'il avait essayées. Il appela ces substances des *lectines,* du mot latin pour « choisir ».

Si une lectine peut choisir entre une sorte de globule rouge et une autre sur la base de différences subtiles de la chimie de surface, on peut trouver des lectines capables de distinguer entre une cellule tumorale et la cellule normale d'origine, qui agglutinent les cellules tumorales et non les cellules normales. Ainsi ces lectines pourraient éliminer les cellules tumorales et ralentir ou inverser la croissance d'un cancer. Des recherches préliminaires ont donné quelques résultats encourageants.

Enfin, plus nous en saurons sur les oncogènes et leur méthode de production, plus nos chances seront grandes de découvrir des moyens d'empêcher leur apparition ou d'encourager leur disparition.

Entre-temps, toutefois, la peur du cancer et le désespoir associé aux traitements entraîne fréquemment le public à se tourner vers des traitements pseudoscientifiques comme ceux attribués à des substances telles que le *krebiozen* et le *laetrile.* Si on ne peut guère blâmer ceux qui se raccrochent à ce qu'ils peuvent, jusqu'à présent ces substances n'ont jamais donné de résultats, et elles ont parfois empêché les patients de suivre des traitements plus efficaces.

Chapitre 15

Le corps

L'alimentation

C'est lorsque les médecins ont compris qu'une bonne santé exigeait un régime simple et équilibré que les premiers grands progrès médicaux ont probablement eu lieu. Les philosophes grecs recommandaient déjà de manger et de boire modérément, non seulement pour des raisons philosophiques, mais aussi parce que ceux qui suivaient ces préceptes se portaient mieux et vivaient plus longtemps. C'était un bon début ; mais par la suite les biologistes s'aperçurent que la modération ne suffisait pas en elle-même. On peut se porter mal même si on a la chance de manger suffisamment et le bon sens de ne pas trop manger, lorsqu'on suit un régime pauvre en denrées essentielles, comme c'est le cas pour beaucoup de gens dans certaines régions du monde.

Les besoins diététiques du corps humain sont assez spécifiques en comparaison avec d'autres organismes. Par exemple, une plante peut se nourrir uniquement à partir de gaz carbonique, d'eau et d'ions inorganiques.

De même, il existe des micro-organismes qui se contentent de n'importe quelle nourriture organique, et qui sont appelés *autotrophes* (« se nourrissant par eux-mêmes »), ce qui signifie qu'ils peuvent pousser dans des environnements où il n'y a pas d'autres êtres vivants. Pour la moisissure du pain *Neurospora,* c'est déjà un peu plus compliqué : elle a besoin de glucides et d'une vitamine appelée biotine en plus de substances inorganiques. Et au fur et à mesure que les formes de vie deviennent de

plus en plus complexes, elles semblent devenir aussi de plus en plus dépendantes pour leur nourriture des éléments organiques de base nécessaires à la fabrication des tissus vivants. Cela tient tout simplement au fait qu'elles ont perdu certaines des enzymes fonctionnant chez les organismes primitifs. Une plante verte dispose d'un ensemble d'enzymes lui permettant de fabriquer à partir de matériaux inorganiques tout ce dont elle a besoin en acides aminés, protéines, lipides et glucides, tandis que *Neurospora* possède toutes ces enzymes, sauf une ou plusieurs de celles qui servent à fabriquer des glucides et de la biotine. Quant aux êtres humains, il leur manque les enzymes nécessaires pour fabriquer de nombreux acides aminés, vitamines et autres produits indispensables qu'il leur faut obtenir déjà préfabriqués dans la nourriture.

Le fait de dépendre de façon croissante de l'environnement peut paraître une sorte de dégénérescence qui désavantage l'organisme, mais ce n'est pas le cas. Si l'environnement fournit les éléments organiques de base, pourquoi faudrait-il que la cellule possède les mécanismes enzymatiques complexes nécessaires à leur fabrication ? En se dispensant de ces mécanismes, la cellule peut utiliser son énergie et son volume à des réactions plus évoluées et plus spécialisées.

Les êtres humains (ainsi que les animaux) doivent ingérer d'autres organismes pour obtenir la nourriture dont ils ont besoin ; ce sont les composés organiques de ces organismes qui constituent ce qu'on appelle la nourriture. Les petites molécules de l'organisme qui est mangé sont absorbées directement par l'intestin de celui qui mange. Les grandes molécules (amidon, protéines, etc.) sont fragmentées ou *digérées* au moyen de systèmes enzymatiques, et les fragments ainsi obtenus (acides aminés, glucose, etc.) peuvent alors être absorbés. Dans le corps de l'organisme qui mange, ces fragments sont soit fragmentés davantage pour produire de l'énergie, soit réassemblés pour former des molécules plus grandes caractéristiques de l'organisme qui mange, et non de celui qui est mangé. De ce point de vue, la vie animale est un cambriolage incessant.

Certains animaux sont *carnivores* et ne mangent que les autres animaux. Mais si tous les animaux étaient carnivores, la vie animale ne durerait pas longtemps, compte tenu du transfert peu efficace d'énergie et de composés organiques entre l'organisme mangé et le mangeur : on peut dire en première approximation qu'il faut dix kilos d'un organisme mangé pour fabriquer un kilo de celui qui mange.

Certains animaux se nourrissent de plantes et sont *herbivores.* La vie végétale étant beaucoup plus répandue que la vie animale, la masse totale des animaux herbivores est beaucoup plus élevée que celle des animaux carnivores. Enfin, certaines espèces animales (dont les êtres humains, les ours et les cochons) sont *omnivores,* et se nourrissent aussi bien de plantes que d'animaux.

Comme le transfert d'énergie et de composés organiques entre les plantes et les animaux qui les mangent est également très peu efficace, la vie s'appauvrirait bientôt jusqu'à ne plus exister si les végétaux ne pouvaient de façon ou d'autre être renouvelés aussi vite qu'ils sont mangés. Ce renouvellement a lieu grâce à l'énergie solaire et au procédé de photosynthèse (voir chapitre 12) qui permet aux végétaux de se nourrir de matières inorganiques, soutenant ainsi pratiquement tout le cycle de la vie, depuis le temps de leur apparition sur la Terre.

La photosynthèse est encore moins efficace que le procédé de digestion des animaux. On estime que moins d'un dixième de pour cent de toute l'énergie solaire qui baigne la Terre est piégée par les plantes et convertie en tissus végétaux ; pourtant, elle suffit à produire entre cent cinquante et deux cents milliards de tonnes de matière organique sèche chaque année. Evidemment, ce procédé ne durera sur la Terre sous sa forme actuelle qu'aussi longtemps que le Soleil restera lui-même sous la même forme – pour encore quelques milliards d'années au moins.

LES CONSTITUANTS ORGANIQUES DE L'ALIMENTATION

Si les aliments ne fournissaient que de l'énergie, nous n'aurions pas vraiment besoin de beaucoup de nourriture. Une demi-livre de beurre suffirait à tous mes besoins en énergie pour une journée, compte tenu de mes occupations sédentaires. Mais la nourriture ne fournit pas que de l'énergie : elle sert aussi de source d'éléments organiques de base destinés à réparer et reconstruire les tissus, qui sont répartis entre toutes sortes d'aliments. A cause de cela, le beurre à lui seul ne suffirait pas à mes besoins.

Le médecin anglais William Prout (celui qui était un siècle en avance sur son temps en suggérant que tous les éléments sont constitués d'hydrogène) fut le premier à suggérer qu'on pouvait diviser les constituants organiques de la nourriture en trois types de substances, appelées par la suite *glucides, lipides* et *protéines.*

Les chimistes et les biologistes du XIX[e] siècle, en particulier l'Allemand Justus von Liebig, s'intéressèrent aux propriétés nutritives de ces substances. Ils découvrirent que les protéines étaient absolument essentielles et suffisaient à l'organisme comme nourriture. En effet, le corps ne peut pas fabriquer de protéines à partir des carbohydrates et des lipides parce que ces substances ne contiennent pas d'azote, mais il peut fabriquer des carbohydrates et des lipides à partir des matériaux contenus dans les protéines. Toutefois, comme les protéines sont relativement rares dans l'environnement, vivre d'un régime entièrement protidique serait du gaspillage – un peu comme de charger un feu avec des meubles quand on dispose de bois coupé pour cela.

Tout le long de l'histoire, et dans bien des endroits encore de nos jours, on a eu à certains moments de la difficulté à trouver suffisamment de nourriture. Soit que la nourriture ne soit pas suffisamment abondante, comme lors d'une famine succédant à une mauvaise récolte ; soit qu'il y ait des défauts dans sa distribution physique ou économique, qui empêchent certaines personnes d'obtenir ou de pouvoir s'acheter de quoi manger.

Même quand la nourriture est suffisante, son contenu en protéines est parfois trop faible, de sorte qu'il y a non pas *sous-nutrition,* mais *malnutrition.* Ces carences en protéines atteignent en particulier les enfants parce qu'ils ont besoin de protéines pour remplacer mais aussi pour construire de nouveaux tissus pendant leur croissance. En Afrique, ces carences en protéines sont particulièrement fréquentes chez les enfants obligés de vivre d'un régime à base de maïs uniquement. Tous les régimes non variés, d'ailleurs, sont dangereux parce que peu d'aliments contiennent tout ce dont on a besoin ; la variété signifie la sécurité.

Par ailleurs, il existe depuis toujours des minorités qui peuvent manger ce qu'elles veulent ; en conséquence, elles ont tendance à manger davantage de chaque aliment qu'elles n'en ont besoin. Ces excès sont stockés par le corps sous forme de lipides (le moyen le plus économique d'empaqueter des calories dans le plus petit espace possible), et grâce à ce phénomène le corps peut stocker des réserves de calories pour les périodes où la nourriture manque. Mais si ces périodes sont inexistantes, la graisse subsiste et la personne devient trop forte, dans les cas extrêmes obèse. Cet état, à la fois désagréable et inconfortable, est souvent associé à de l'artériosclérose ou à d'autres maladies ; l'excès de poids ne garantit pas non plus contre les carences en éléments essentiels quand le régime n'est pas correctement équilibré.

Dans un pays comme les États-Unis où le niveau de vie est particulièrement élevé et où la minceur est un critère de beauté, l'excès de poids est un réel problème. Le seul moyen raisonnable de l'éviter est de diminuer la quantité de nourriture ingérée ou d'augmenter l'activité (ou les deux ensemble), et les gens qui se refusent à faire l'un ou l'autre sont condamnés à être trop forts, indépendamment des astuces auxquelles ils ont recours.

LES PROTÉINES

Dans l'ensemble, les aliments riches en protéines ont tendance à être plus coûteux et plus rares (ces deux caractéristiques vont en général de pair) que ceux qui sont pauvres en protéines ; et la teneur en protéines est plus élevée dans les nourritures animales que dans les nourritures végétales.

Il y a là un problème pour ceux qui choisissent d'être végétariens, car s'il est vrai que le régime végétarien a ses avantages, ceux qui le pratiquent doivent essayer de s'assurer une ration suffisante de protéines ; c'est possible, puisqu'il suffit de soixante grammes de protéines par jour en moyenne pour nourrir un adulte, et d'un peu plus pour les enfants, les femmes enceintes et les mères qui allaitent.

Tout dépend bien sûr des protéines choisies. Au XIXe siècle des expériences ont été faites pour déterminer si une population pouvait se nourrir en période de famine avec de la *gélatine,* une matière protéique que l'on obtient en chauffant les os, les tendons et autres parties des animaux qui ne peuvent être mangées autrement. Le physiologiste français François Magendie a démontré qu'un chien perd du poids et meurt quand il prend de la gélatine comme seule nourriture. Non qu'il y ait quelque chose de dangereux à utiliser de la gélatine comme source de nourriture, mais simplement elle ne fournit pas tous les matériaux organiques nécessaires en l'absence d'autres protéines dans le régime. Encore une fois, c'est dans la variété que se situe la sécurité.

La valeur nutritive d'une protéine dépend de l'efficacité avec laquelle le corps peut utiliser l'azote qu'elle fournit. Les premières expériences sur l'*équilibre de l'azote* ont eu lieu en 1854 : les agronomes John Bennet Lawes et Joseph Henry Gilbert, qui avaient donné à manger à des cochons des protéines sous deux formes différentes (lentilles et orge), se rendirent compte que les cochons fixaient beaucoup mieux l'azote de l'orge que celui des lentilles.

Un organisme qui grandit normalement accumule progressivement l'azote de la nourriture qu'il ingère *(équilibre positif de l'azote)*. Mais lorsqu'il est en état de sous-alimentation, son corps souffre de consomption, et si la seule source de protéines est par exemple la gélatine, il continue à être sous-alimenté et à dépérir, du point de vue de l'équilibre de l'azote (une situation appelée *équilibre négatif de l'azote*).

Il perd plus d'azote qu'il n'en absorbe, indépendamment de la quantité de gélatine qu'il absorbe.

Pourquoi en est-il ainsi ? Les chimistes du XIXe siècle ont découvert finalement que la gélatine est une protéine particulièrement simple, qui ne contient pas de tryptophane ni certains des acides aminés présents dans la plupart des protéines. Sans ces matériaux de construction, le corps ne peut pas fabriquer les protéines dont il a besoin pour sa propre substance. Par conséquent, sauf s'il obtient d'autres protéines dans sa nourriture en même temps, les acides aminés qui se trouvent dans la gélatine sont inutiles et doivent être excrétés. C'est comme si les constructeurs d'une maison se retrouvaient avec beaucoup de bois de charpente mais pas de clous. Non seulement ils ne pourraient pas construire la maison, mais le bois serait en trop et il faudrait s'en débarrasser. Des essais ont été faits en 1890 – mais sans succès – pour rendre la gélatine plus efficace comme aliment de régime, en lui additionnant certains des acides aminés qui lui font défaut. De meilleurs résultats ont été obtenus avec des protéines plus complexes que la gélatine.

Par la suite, en 1906, les biochimistes anglais Frederick Gowland Hopkins et Edith Gertrude Willcock donnèrent à manger à des souris un régime dans lequel la *zéine* était la seule protéine (une protéine que l'on trouve dans le maïs, et qui est très pauvre en acide aminé tryptophane). Les souris en moururent au bout d'environ quatorze jours (le manque de tryptophane est la cause d'une maladie par carence appelée *Kwashiorkor,* fréquente chez les enfants africains). Les chercheurs essayèrent ensuite de nourrir des souris avec de la zéine et du tryptophane, et cette fois les souris survécurent deux fois plus longtemps. C'était la première preuve véritable que les acides aminés, plus que les protéines, sont les composants essentiels d'un régime. La mort prématurée des souris était probablement due au manque de certaines vitamines inconnues à cette époque.

En 1930, le diététicien américain William Cumming Rose se pencha sérieusement sur le problème des acides aminés. On connaissait alors les principales vitamines de sorte qu'il put les fournir aux animaux et porter son attention sur les acides aminés. Rose donna d'abord à manger à des rats un mélange d'acides aminés plutôt que de protéines, mais ils ne vécurent pas longtemps de ce régime. Pourtant, lorsqu'il leur donnait à manger de la caséine, une protéine présente dans le lait, ils grandissaient normalement. Il y avait donc dans la caséine quelque chose – selon toute probabilité, un acide aminé inconnu – qui n'était pas présent dans le mélange d'acides aminés qu'il utilisait. Il décomposa la caséine en ses éléments de base et essaya d'ajouter les divers fragments moléculaires à son mélange d'acides aminés. C'est ainsi qu'il trouva l'acide aminé *thréonine,* le dernier des principaux acides aminés à avoir été découvert. Quand il ajouta la thréonine extraite de la caséine à son mélange d'acides aminés, les rats grandirent normalement, sans qu'il y ait de protéine intacte dans leur régime.

Rose enleva alors les acides aminés les uns après les autres du régime et identifia de cette façon dix acides aminés indispensables au régime du

rat : lysine, tryptophane, histidine, phenylalanine, leucine, isoleucine, thréonine, méthionine, valine et arginine. Lorsqu'on les fournit en grosses quantités à des rats, ceux-ci peuvent fabriquer tous les autres acides aminés dont ils ont besoin (par exemple glycine, proline, acide aspartique, alanine et ainsi de suite).

En 1940, Rose se tourna vers la question des exigences humaines en acides aminés. Ayant persuadé ses étudiants en doctorat de se soumettre à un régime surveillé dans lequel un mélange d'acides aminés constituait la seule source d'azote, il put en 1949 annoncer que l'homme adulte a besoin seulement de huit acides aminés dans son régime : phenylalanine, leucine, isoleucine, méthionine, valine, lysine, tryptophane et thréonine. Il semblerait que les êtres humains soient moins spécialisés que le rat de ce point de vue, puisque l'arginine et l'histidine indispensables chez le rat ne le sont pas chez l'homme (ou n'importe quel autre mammifère pour lequel l'expérience a été faite).

Théoriquement une personne peut se contenter des huit acides aminés essentiels comme régime, si elle en a suffisamment, et peut fabriquer à partir de là non seulement tous les autres acides aminés dont elle a besoin, mais aussi des carbohydrates et des lipides. En fait, un régime composé seulement d'acides aminés est coûteux, sans tenir compte de sa monotonie et de son manque de saveur. Mais il est fort utile d'avoir une image complète de nos besoins en acides aminés pour pouvoir éventuellement compléter les protéines naturelles et obtenir une efficacité maximale d'absorption et d'utilisation de l'azote.

LES LIPIDES

Les lipides, eux aussi, peuvent être décomposés en éléments de base plus simples, dont les principaux sont les *acides gras.* Ceux-ci sont *saturés* si leur molécule comporte tous les atomes d'hydrogène qu'elle peut contenir, et *non saturés* si une paire ou plus d'atomes d'hydrogène leur manque ; si plus d'une paire d'atomes d'hydrogène est absente, ils sont dits *polyinsaturés.*

Les lipides qui contiennent des acides gras non saturés ont tendance à fondre à des températures plus basses que ceux qui contiennent des acides gras saturés. Comme il est préférable d'avoir des lipides sous forme liquide dans l'organisme, les plantes et les animaux à sang froid contiennent souvent des lipides plus insaturés que les oiseaux et les mammifères qui ont un sang chaud. Le corps humain, lui, ne peut pas fabriquer des lipides polyinsaturés à partir de lipides saturés et contient donc des acides gras polyinsaturés appelés *acides gras essentiels.* C'est l'un des avantages des régimes végétariens, qui risquent moins d'engendrer des carences en acides gras.

Les vitamines

Le grand public se laisse trop facilement induire en erreur par les modes en matière d'alimentation et par les best-sellers qui promettent n'importe

quelle guérison – même en notre époque de progrès. En fait, c'est peut-être parce que notre époque est celle du progrès que les modes alimentaires sont possibles. Dans la plus grande partie de l'histoire humaine, la nourriture a consisté essentiellement en ce qui pouvait être produit sur place, et qui n'était pas très abondant. Si on ne mangeait pas ce qu'on avait à sa disposition, on mourait de faim ; on ne pouvait se permettre de faire le difficile, et les modes alimentaires n'existaient pas.

Grâce aux transports modernes, l'envoi de nourriture d'une partie de la Terre à l'autre est devenu possible, en particulier avec l'usage de la réfrigération à grande échelle, et les menaces de famine ont ainsi diminué. Auparavant, les famines étaient toujours locales ; les provinces environnantes pouvaient regorger de nourriture sans que celle-ci puisse être transportée vers les zones où sévissait la famine.

L'accumulation de réserves de nourriture chez soi est devenue possible quand on apprit à conserver les aliments en les séchant, en les salant, en augmentant leur contenu en sucre, en les faisant fermenter et ainsi de suite. Il est devenu possible de conserver les aliments dans un état proche de celui d'origine grâce aux méthodes de stockage sous vide après cuisson (le fait de cuire la nourriture tue les micro-organismes, et le stockage sous vide empêche ceux qui ont résisté de croître et de se reproduire). Le stockage sous vide a été utilisé pour la première fois par un chef cuisinier français, François Appert, qui développa cette technique lors d'un concours organisé par Napoléon Ier, destiné à trouver le meilleur mode de conservation des aliments pour les armées. Appert faisait usage de récipients en verre ; mais aujourd'hui, on utilise des boîtes recouvertes de fer-blanc, ou *boîtes de conserves.* Depuis la Seconde Guerre mondiale, les aliments frais surgelés se sont généralisés, et le nombre croissant de congélateurs dans les maisons a encore amélioré la distribution et la variété des aliments frais. Chaque amélioration de la distribution en aliments a augmenté la possibilité de pratiquer les modes alimentaires.

LES MALADIES PAR CARENCE

Je ne dis pas qu'un choix intelligent de la nourriture ne soit pas nécessaire. Dans certains cas, des aliments spécifiques peuvent guérir efficacement une maladie. C'est le cas des *maladies par carence* qui apparaissent alors que les aliments ainsi que les protéines sont présents en quantité suffisante. Elles sont produites par l'absence dans le régime d'une substance dont la présence en petite quantité est essentielle au fonctionnement chimique du corps. Ces maladies apparaissent presque toujours chez des gens qui ne bénéficient pas d'un régime normal et équilibré – c'est-à-dire contenant toute une variété d'aliments différents.

Il est vrai que la valeur d'un régime équilibré et varié était déjà comprise par beaucoup de médecins au XIXe siècle ou même avant, alors que la chimie de la nourriture restait un mystère. Un exemple célèbre est celui de Florence Nightingale, une infirmière anglaise héroïque qui, pendant la guerre de Crimée, avait réussi à nourrir correctement les soldats et à leur procurer des soins médicaux convenables. Mais la *diététique* (l'étude systématique du régime alimentaire) dut attendre la fin du siècle, et la découverte de certaines susbtances essentielles à la vie dans la nourriture quoique présentes à l'état de traces.

Autrefois on connaissait déjà le *scorbut*, une maladie qui se caractérise par les symptômes suivants : les capillaires deviennent de plus en plus fragiles, les gencives saignent, les dents tombent, les blessures se guérissent avec difficulté quand elles le font, le patient devient faible et finit par mourir. Cette maladie se répandait surtout dans les villes assiégées et pendant les longs voyages sur l'Océan (on la signale à bord du bateau de Vasco de Gama pendant son voyage autour de l'Afrique en direction de l'Inde en 1497, et l'équipage de Magellan souffrit plus du scorbut que de la sous-alimentation pendant son premier voyage autour du monde, une génération plus tard). Les bateaux qui partaient pour de longs voyages, à cette époque où la réfrigération n'existait pas, ne pouvaient emporter que de la nourriture qui ne s'abîme pas, autrement dit des biscuits de mer et du porc salé. Mais les médecins ne firent pas le lien entre le scorbut et le régime alimentaire pendant plusieurs siècles.

En 1536, au cours de l'hiver que l'explorateur français Jacques Cartier passait au Canada, cent dix de ses hommes furent atteints par le scorbut. Les indigènes suggérèrent un remède qu'ils connaissaient : boire de l'eau dans laquelle des aiguilles de pin avaient trempé, et pour avoir suivi ce conseil pourtant fort simple, les hommes de Jacques Cartier guérirent de leur scorbut.

Deux siècles plus tard, en 1747, le médecin écossais James Lind essaya d'utiliser comme remède des fruits frais et des légumes. Il s'aperçut, en traitant des marins atteints du scorbut, que les oranges et les citrons amenaient la guérison rapidement. Et le capitaine Cook, en voyage d'exploration à travers le Pacifique entre 1772 et 1775, protégea son équipage du scorbut en l'obligeant à manger régulièrement de la choucroute. Néanmoins, ce n'est qu'en 1795 que les officiers de la marine anglaise furent suffisamment impressionnés par l'expérience de Lind (et suffisamment convaincus qu'une flotte atteinte du scorbut pouvait perdre une bataille navale presque sans se battre) pour imposer aux marins britanniques des rations journalières de jus de citron. Grâce au jus de citron le scorbut a disparu de la marine anglaise.

De même, un siècle plus tard, en 1884, l'amiral Kanehiro Takaki de la marine japonaise remplaça sur ses bateaux le régime à base de riz seul par un régime plus diversifié. C'est ainsi que la maladie appelée béribéri prit fin dans la marine japonaise.

En dépit de quelques victoires de cette sorte sur les régimes alimentaires (que personne ne pouvait expliquer), les biologistes du XIXe siècle refusaient de croire qu'une maladie puisse guérir pour des raisons alimentaires, en particulier au moment où la théorie microbienne des maladies faisait son apparition. Toutefois, en 1896, le médecin hollandais Christiaan Eijkman le démontra presque malgré lui.

Eijkman fut envoyé dans les Indes néerlandaises pour faire des recherches sur le béribéri, qui était endémique dans cette région (et qui encore aujourd'hui, alors que l'on connaît ses causes et ses remèdes, tue cent mille personnes par an). Takaki avait réussi à enrayer le béribéri par des mesures diététiques, mais l'Occident, apparemment, n'avait pas confiance en ce qui ressemblait à une science mystique de l'Orient.

Pensant que le béribéri était d'origine bactérienne, Eijkman emporta avec lui des poulets comme animaux d'expérience pour mettre en évidence le microbe concerné. Un heureux hasard déjoua ses plans. Brusquement, tous ses poulets tombèrent malades d'une maladie qui les rendait

paralytiques et certains en moururent, tandis que d'autres retrouvèrent la santé après environ quatre mois. Eijkman, intrigué de ne pas avoir trouvé le microbe responsable de la maladie, orienta finalement ses recherches vers le régime alimentaire des poulets. Il découvrit que la personne responsable de nourrir les poulets avait, pour faire des économies (et elle en avait sûrement profité), utilisé des restes de nourriture, en particulier du riz blanc en provenance de l'hôpital militaire. Au bout de quelques mois, un nouveau cuisinier était arrivé, qui lui aussi nourrissait les poulets ; il avait mis fin à cette corruption mesquine en fournissant aux poulets leur nourriture habituelle, du riz non décortiqué. C'est ainsi que les poulets avaient guéri.

Eijkman réalisa alors une expérience. Il mit les poulets à un régime de riz blanc qui les fit tomber malades. Une fois nourris avec du riz non décortiqué, ils se rétablirent. C'était le premier cas de maladie par carence alimentaire produit artificiellement. Eijkman décida que cette polynévrite qui atteignait les poulets ressemblait par ses symptômes au béribéri de l'homme. Peut-être les êtres humains attrapaient-ils le béribéri parce qu'ils mangeaient seulement du riz blanc.

Le riz utilisé pour la consommation humaine était débarrassé de ses cosses essentiellement pour qu'il se conserve mieux, parce que le germe du riz enlevé avec les cosses contient des lipides qui deviennent facilement rances. Eijkman et son collaborateur Gerrit Grijns décidèrent de mettre en évidence ce qui, dans les cosses du riz, protégeait du béribéri. Ils réussirent à dissoudre le facteur responsable dans de l'eau, et s'aperçurent que celui-ci traversait des membranes imperméables aux protéines ; cette substance ne pouvait donc être qu'une très petite molécule. Ils ne purent toutefois pas l'identifier.

Pendant ce temps d'autres chercheurs découvraient des facteurs mystérieux qui semblaient essentiels à la vie. En 1905 un diététicien hollandais, Cornelis Adrianis Pekelharing, observa que toutes ses souris, lorsqu'elles étaient nourries d'un régime artificiel pourtant riche en lipides, protéines, et carbohydrates, mouraient dans le mois, mais qu'elles se portaient bien quand il ajoutait quelques gouttes de lait à ce régime. En Angleterre, le biochimiste Frederick Hopkins, qui avait déjà démontré l'importance des acides aminés dans le régime alimentaire, entreprit une série d'expériences dans lesquelles lui aussi démontra qu'un facteur de la caséine du lait permettait la croissance lorsqu'il était ajouté à un régime alimentaire artificiel. Cette substance était soluble dans l'eau. Encore mieux que la caséine comme complément au régime, il y avait l'extrait de levures en petite quantité.

Eijkman et Hopkins reçurent le prix Nobel de médecine et de physiologie en 1929 pour leur travail innovateur dans l'identification des substances essentielles à la vie à l'état de traces dans le régime alimentaire.

L'IDENTIFICATION DES VITAMINES

Il fallait maintenant isoler ces facteurs vitaux présents à l'état de traces dans la nourriture. En 1912 trois biochimistes japonais, Umetaro Suzuki, T. Shimamura et S. Ohdake, parvinrent à extraire des cosses de riz un produit très efficace pour lutter contre le béribéri. Il suffisait de cinq à dix milligrammes de ce produit pour guérir des poulets de cette maladie. La même année,

le biochimiste d'origine polonaise Casimir Funk (qui travaillait en Angleterre et devait plus tard aller aux États-Unis) réussit à préparer ce même composé en partant de la levure.

Il s'avéra que ce composé était une amine (c'est-à-dire une molécule contenant le groupe amine NH_2) et Funk l'appela *vitamine*, le mot latin pour « amine vitale ». Il émit l'hypothèse que le béribéri, la pellagre et le rachitisme provenaient tous de déficiences en vitamines. L'hypothèse de Funk était correcte en ce sens que ces maladies sont produites par des carences alimentaires. Mais les « vitamines » n'étaient pas toutes des amines.

En 1913, deux biochimistes américains, Elmer Vernon McCollum et Marguerite Davis, découvrirent un autre facteur essentiel à la vie présent à l'état de traces dans le beurre et le jaune d'œuf. Celui-ci était soluble dans les corps gras et non dans l'eau. McCollum l'appela *facteur A liposoluble*, par opposition au *facteur B hydrosoluble*, nom qu'il donna au facteur qui guérissait le béribéri. C'est ainsi que commença la coutume d'appeler ces facteurs par des lettres. En 1920, le biochimiste anglais Jack Cecil Drummond les rebaptisa *vitamine A* et *vitamine B*. Il suggéra aussi que le facteur qui guérissait le scorbut était une troisième substance de la même sorte, qu'il appela *vitamine C*.

La vitamine A fut rapidement identifiée comme facteur alimentaire indispensable pour empêcher le développement d'une sécheresse anormale des muqueuses qui entourent l'œil, la xérophtalmie, du mot grec signifiant « yeux secs ». En 1920, McCollum et ses collaborateurs tirèrent de l'huile de foie de morue une substance qui était capable de guérir à la fois la xérophtalmie et la maladie des os appelée *rachitisme*, et qu'on pouvait modifier de sorte qu'elle guérisse uniquement le rachistime. Ils décidèrent que le facteur antirachitique était une quatrième vitamine qu'ils appelèrent *vitamine D*. Les vitamines D et A sont liposolubles, tandis que les vitamines C et B sont hydrosolubles.

En 1930, il était devenu évident que la vitamine B n'était pas une substance simple, mais un mélange de produits chimiques ayant des propriétés différentes. Le facteur alimentaire qui guérissait le béribéri fut appelé *vitamine* B_1, un deuxième facteur appelé *vitamine* B_2 et ainsi de suite. Il y eut de nouveaux facteurs annoncés qui s'avérèrent de fausses nouvelles, de sorte qu'on n'entend par exemple plus parler de vitamine B_3, B_4 ni B_5. Toutefois, on est allé jusqu'à la vitamine B_{14}. L'ensemble de ces vitamines (toutes hydrosolubles) est fréquemment appelé *complexe de vitamines B*.

De nouvelles lettres sont également apparues. Les *vitamines E* et *K* (toutes deux liposolubles) sont de véritables vitamines, tandis que la *vitamine F* n'est pas une vitamine et que la *vitamine H* appartient au complexe B.

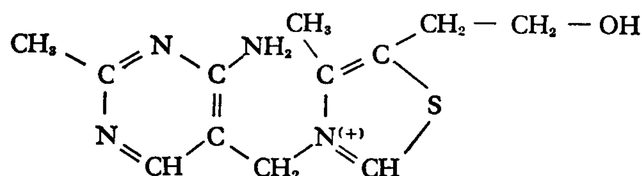
A présent que leur composition chimique a été élucidée, on n'utilise plus guère les lettres des vraies vitamines, qui sont plus connues sous leur nom chimique ; néanmoins, les vitamines liposolubles ont gardé plus souvent leur désignation par lettre que les hydrosolubles.

COMPOSITION CHIMIQUE ET STRUCTURE

Il n'a pas été facile d'établir la composition chimique et la structure de ces vitamines parce qu'elles n'apparaissent qu'en très petite quantité. Pour

donner un exemple, une tonne de cosses de riz contient seulement environ cinq grammes de vitamine B_1. Ce n'est qu'en 1926 qu'on a réussi à extraire cette vitamine en quantité suffisante et suffisamment pure pour pouvoir l'analyser chimiquement. La composition de la vitamine B établie par les deux biochimistes hollandais Barend Coenraad Petrus Jansen et William Frederick Donath à partir d'un petit échantillon se révéla fausse. En 1932, Ohdake obtint pratiquement la bonne formule en essayant de nouveau sur un échantillon un peu plus grand. Il fut le premier à détecter un atome de soufre dans une molécule de vitamine.

Finalement, en 1934, Robert Runnels Williams, alors directeur du département de chimie aux Laboratoires Bell, vit culminer vingt ans de recherche lorsqu'après avoir laborieusement séparé la vitamine B_1 à partir de plusieurs tonnes de cosses de riz, il réussit à en établir la formule complète, qui est la suivante :

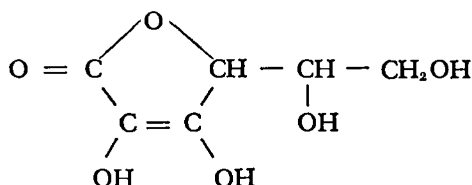


Parce que le trait le plus caractéristique de la molécule est son atome de soufre (*theion* en grec), cette vitamine a été appelée *thiamine*.

Pour la vitamine C, le problème est différent. Les agrumes en sont une source relativement riche, mais la difficulté était de trouver un animal de laboratoire qui ne fabrique pas sa propre vitamine C. La plupart des mammifères, mis à part les êtres humains et autres primates, ont conservé la capacité de fabriquer cette vitamine. Sans un animal expérimental simple et peu coûteux qui puisse être atteint par le scorbut, il était difficile de suivre la localisation de la vitamine C dans les diverses fractions obtenues à partir des jus de fruits au cours du processus chimique.

C'est en 1918 que les biochimistes américains B. Cohen et Lafayette Benedict Mendel s'aperçurent que les cobayes ne pouvaient pas fabriquer cette vitamine. En fait, ils sont atteints par le scorbut bien plus facilement que les hommes. Mais il subsistait une autre difficulté. La vitamine C, très instable (c'est la plus instable des vitamines), se perdait facilement au cours du processus chimique d'isolement. Un certain nombre de chercheurs se mirent sans succès en quête de cette vitamine.

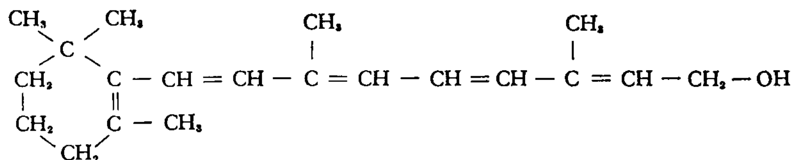
La vitamine C fut finalement isolée par quelqu'un qui n'en avait pas l'intention. En 1928, le biochimiste d'origine hongroise Albert Szent-Györgi travaillait à Londres dans le laboratoire de Hopkins avec pour principal sujet d'étude l'utilisation de l'oxygène par les tissus ; c'est ainsi qu'il isola une substance contenue dans les choux, qui permettait le transfert des atomes d'hydrogène d'un composé à un autre. Peu de temps après, Charles Glen King et ses collaborateurs à l'université de Pittsburgh qui, eux, cherchaient vraiment à isoler la vitamine C, firent une préparation de la substance présente dans les choux et découvrirent qu'elle protégeait bien contre le scorbut. De plus ils remarquèrent que c'était la même substance que celle de cristaux obtenus à partir de jus de citron. King en détermina la structure en 1933, à savoir une molécule de sucre à six atomes de carbone, appartenant à la série L au lieu de la série D :



On l'appela *acide ascorbique*, du grec signifiant « pas de scorbut ».

Quant à la structure de la vitamine A, on observa que les nourritures riches en vitamine A (le beurre, le jaune d'œuf, les carottes, l'huile de foie de poisson, etc.) sont souvent jaunes ou orangées. La substance à l'origine de cette couleur est un hydrocarbure appelé *carotène* ; et en 1929, le biochimiste anglais Thomas Moore démontra que les rats nourris d'un régime contenant du carotène stockaient de la vitamine A dans leur foie. La vitamine elle-même n'est pas de couleur jaune ; par conséquent, bien que le carotène ne soit pas lui-même de la vitamine A, c'est le foie qui le convertit en vitamine A (le carotène est maintenant considéré comme un exemple de provitamine).

En 1937, les chimistes américains Harry Nicholls Holmes et Ruth Elisabeth Corbet isolèrent la vitamine A sous forme cristallisée à partir de l'huile de foie de poisson. C'est un composé à vingt atomes de carbone – une moitié de la molécule de carotène plus un groupe hydroxyle :

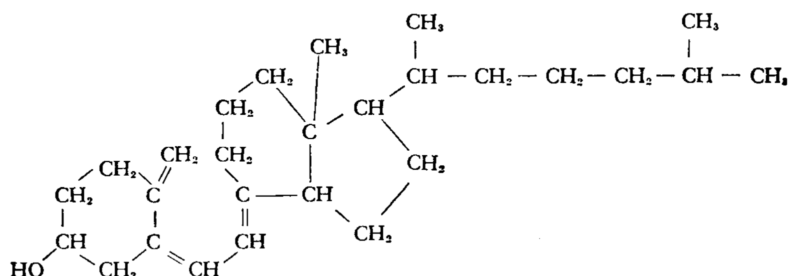


C'est grâce à la lumière du soleil que les chimistes, ayant pour objectif d'isoler la vitamine D, firent des progrès. Dès 1921, le groupe de McCollum (le premier à avoir démontré l'existence de la vitamine) observa que les rats ne développent pas de rachitisme, quand ils sont nourris sans vitamine D, s'ils sont exposés à la lumière du soleil. Les biochimistes en conclurent que l'énergie de la lumière solaire convertit une certaine provitamine du corps en vitamine D. Comme la vitamine D est liposoluble, c'est dans les lipides alimentaires qu'ils cherchèrent la provitamine.

En fractionnant les lipides et en exposant chaque fraction séparément à la lumière solaire, ils démontrèrent que la provitamine convertie en vitamine D par la lumière est un stéroïde. Lequel ? Ce n'était pas le cholestérol, qui est le stéroïde le plus fréquent du corps. En 1926, les biochimistes anglais Otto Rosenheim et T.A. Webster découvrirent que la lumière solaire convertissait un stérol apparenté, l'ergostérol (ainsi appelé du fait qu'il a été isolé à partir de l'ergot de seigle), en vitamine D. Le chimiste allemand Adolf Windaus fit cette découverte indépendamment à peu près au même moment. Il reçut le prix Nobel de chimie en 1928 pour ce travail et ses autres travaux sur les stéroïdes.

On constate toutefois que l'ergostérol n'apparaît pas chez les animaux. C'est le *7-déhydrocholestérol* qui est réellement la provitamine humaine ; il

diffère du cholestérol seulement par deux atomes d'hydrogène en moins. La vitamine D présente alors la formule suivante :



L'une des formes de la vitamine D s'appelle le *calciférol*, du latin pour « transport de calcium », parce qu'il est essentiel à la formation de la structure osseuse.

Les carences en vitamines n'aboutissent pas toujours à l'apparition d'une maladie grave. En 1922, Herbert McLean Evans et K. J. Scott, à l'université de Californie, ont pensé que le manque d'une vitamine était à l'origine de la stérilité chez les animaux. C'est seulement en 1936 que Evans et son groupe réussirent à isoler celle-ci, la vitamine E. On l'appela alors *tocophérol* (du mot grec pour « porter les enfants »).

On ignore malheureusement si les êtres humains ont besoin de vitamine E, et de combien. Bien sûr, il est hors de question de pratiquer des expériences de régime alimentaire sur des êtres humains dans le but d'induire la stérilité. Même chez l'animal, le fait qu'on puisse le rendre stérile en ne lui donnant pas de vitamine E ne signifie pas nécessairement que la stérilité apparaît ainsi.

Vers 1930, le biochimiste danois Carl Peter Henrik Dam découvrit empiriquement sur des poulets l'existence d'une vitamine impliquée dans la coagulation sanguine. Il l'appela *vitamine de coagulation*, qui devint par la suite la *vitamine K*. Edward Doisy et ses collaborateurs, à l'université de Saint Louis, isolèrent par la suite la vitamine K et déterminèrent sa structure. Dam et Doisy partagèrent le prix Nobel de médecine et de physiologie en 1943.

La vitamine K ne fait pas partie des vitamines principales et ne constitue pas un problème alimentaire. En temps normal, une quantité plus que suffisante de cette vitamine est fabriquée par les bactéries de l'intestin. En fait, celles-ci en font tellement que les défécations contiennent parfois plus de vitamine K que la nourriture. Ce sont surtout les nouveau-nés qui sont menacés par le danger d'une mauvaise coagulation et d'hémorragies dues à une déficience en vitamine K. Dans les conditions d'hygiène des hôpitaux modernes, il faut aux nourrissons trois jours pour accumuler une quantité raisonnable de bactéries intestinales et ils sont protégés par des injections de vitamine K, soit sur eux directement soit sur leur mère juste avant la naissance. Autrefois, les nourrissons accumulaient ces bactéries pratiquement dès la naissance, et si les enfants mouraient de maladies infectieuses ou autres, ils étaient protégés contre les dangers d'hémorragies.

En fait, on peut se demander si l'on pourrait vivre sans ces bactéries intestinales ou, au contraire, si cette symbiose n'est pas si étroite qu'elle en est devenue indispensable. Des animaux (par exemple des souris), dépourvus de microbes et élevés dès la naissance dans des conditions

complètement stériles, peuvent se reproduire pendant douze générations de cette façon. De telles expériences ont été faites en 1928.

A la fin des années trente et au début des années quarante, les biochimistes ont identifié plusieurs autres vitamines B qui, maintenant, portent les noms de *biotine, acide pantothénique, pyridoxine, acide folique* et *cyanocobalamine*. Elles sont toutes fabriquées par les bactéries intestinales ; de plus, elles sont présentes dans tous les aliments ; on n'a noté aucun cas de carence alimentaire. En fait, les chercheurs ont dû nourrir des animaux artificiellement en les excluant volontairement, et même en ajoutant des antivitamines qui neutralisent celles des bactéries intestinales, pour mettre en évidence les symptômes de déficience (les antivitamines sont des substances semblables aux vitamines par leur structure, qui immobilisent l'enzyme à laquelle sert la vitamine par inhibition compétitive).

LES THÉRAPIES VITAMINÉES

La détermination de la structure des différentes vitamines a en général été très vite suivie (parfois même précédée) de leur synthèse. La thiamine, par exemple, a été synthétisée par Williams et son groupe trois ans après que ceux-ci eurent trouvé sa structure ; l'acide ascorbique, un peu avant que sa structure soit complètement élucidée par King, a été synthétisé par le biochimiste suisse d'origine polonaise, Tadeus Reichstein et son groupe, en 1933. La vitamine A, pour prendre un autre exemple, a été synthétisée en 1936 par deux groupes différents de chimistes (là aussi avant que sa structure soit complètement élucidée).

L'usage des vitamines synthétiques a rendu possibles l'addition de fortifiants à la nourriture (en commençant par le lait, dès 1924), et la préparation de mélanges de vitamines à des prix raisonnables vendus en pharmacie. Les besoins en vitamines sont variables selon les individus. De toutes les vitamines, celle qui a tendance à faire le plus défaut dans l'alimentation est la vitamine D. Les jeunes enfants des pays nordiques, où la lumière du soleil est faible en hiver, risquent de devenir rachitiques et peuvent avoir besoin de suppléments vitaminés ou d'aliments irradiés. Mais la dose de vitamine D (et de vitamine A) doit être surveillée attentivement, parce qu'un excès de ces vitamines peut être nocif. Pour ce qui est des vitamines B, toute personne ayant un régime alimentaire normal n'a pas besoin de médicaments. De même pour la vitamine C, qui de toute façon ne devrait pas poser de problèmes, car rares sont ceux qui n'aiment pas le jus d'orange ou qui n'en boivent pas régulièrement, maintenant que les vitamines sont connues.

Dans l'ensemble, l'usage généralisé des vitamines, même s'il contribue principalement au profit des industries pharmaceutiques, ne fait pas de mal, et il explique peut-être pourquoi la génération actuelle d'Américains est plus grande et plus forte que les générations précédentes.

Au cours des années soixante-dix, on a proposé des *thérapies mégavitaminées*. On pensait en effet que les quantités minimales de vitamines nécessaires à la prévention des maladies par carence ne suffisaient pas obligatoirement au fonctionnement optimal du corps ou à la prévention d'autres maladies. On prétendait par exemple que d'importantes doses de vitamine B amélioreraient l'état des schizophrènes.

Le défenseur le plus célèbre des thérapies mégavitaminées est Linus Pauling qui, en 1970, a soutenu que de fortes doses quotidiennes de vita-

mine C prévenaient les rhumes et avaient des effets bénéfiques sur la santé. Il n'a pas réussi à convaincre le corps médical dans son entier, mais le grand public, qui croit d'autant plus aux vitamines qu'elles sont faciles à obtenir et bon marché, a vidé les stocks de vitamine C des pharmaciens.

Des excès de vitamines hydrosolubles comme le complexe B et la vitamine C ne peuvent pas faire de mal, car elles ne restent pas dans le corps ; elles sont facilement excrétées. Par conséquent lorsqu'une dose élevée n'a pas d'usage spécifique dans le corps, l'excès qui en résulte sert tout simplement à enrichir l'urine.

Le cas est différent pour les vitamines liposolubles, en particulier A et D. Elles ont tendance à se dissoudre dans la graisse du corps et à y être stockées ; elles y restent relativement immobiles, comme la graisse elle-même. Un excès peut alors représenter une surcharge pour le corps, le rendre moins fonctionnel, et donner lieu à des *hypervitaminoses*. La vitamine A est stockée par le foie, en particulier chez les poissons et les animaux qui en mangent (ce qui fait qu'on a imposé à toute une génération de jeunes enfants des doses régulières d'*huile de foie de morue*) ; et on raconte des histoires horribles d'explorateurs de l'Arctique tombés malades ou même morts pour s'être nourris de foies d'ours polaires – empoisonnés par la vitamine A.

LES VITAMINES EN TANT QU'ENZYMES

Les biochimistes cherchèrent bien sûr à comprendre comment les vitamines, présentes dans le corps en si petite quantité, peuvent exercer des effets si importants sur la chimie du corps. Peut-être était-ce dû à leur association aux enzymes, qui elles aussi agissent en petite quantité.

La réponse fut apportée par des études détaillées sur la chimie des enzymes. Les chimistes des protéines savaient depuis longtemps que certaines protéines ne sont pas constituées uniquement d'acides aminés, et possèdent des groupes prosthétiques de natures différentes, comme par exemple le groupe hème de l'hémoglobine (voir chapitre 11). Ces groupes prosthétiques ont en général tendance à être fortement attachés au reste de la molécule. Toutefois, il existe certains cas d'enzymes pour lesquels les parties non protéiques sont attachées assez faiblement et s'enlèvent facilement.

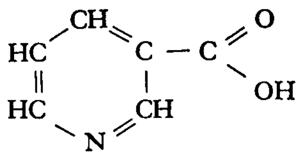
Tout cela fut découvert en 1904 par Arthur Harden (qui devait bientôt découvrir les composés intermédiaires à base de phosphore, voir chapitre 12). Il travaillait sur un extrait de levures destiné à provoquer la fermentation des sucres. Après l'avoir placé dans un sac constitué d'une membrane semi-perméable, il plaça le sac dans de l'eau douce de sorte que les petites molécules pouvaient traverser la membrane, mais les grandes molécules protéiques ne le pouvaient pas. Il laissa la dialyse progresser, et au bout d'un moment, il découvrit que l'activité de l'extrait avait disparu. Ni le liquide à l'intérieur, ni celui à l'extérieur du sac ne causait la fermentation du sucre. Lorsque les deux liquides étaient mélangés, l'activité réapparaissait.

L'enzyme était donc constituée non seulement d'une grande molécule de protéine, mais aussi d'une molécule de *coenzyme*, suffisamment petite pour traverser les pores de la membrane. Le coenzyme était essentiel à l'activité enzymatique (dont il est la partie active, pour ainsi dire).

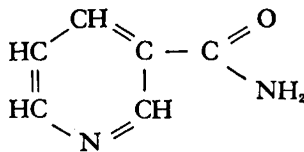
Les chimistes se sont tout de suite attaqués au problème de la détermination de la structure du coenzyme (et d'autres adjonctions semblables sur d'autres enzymes). Le chimiste suédois d'origine allemande Hans Karl August Simon

von Euler-Chelpin a été le premier à faire des progrès réels de ce point de vue, et a obtenu pour ce travail le prix Nobel de chimie en 1929.

Le coenzyme de l'enzyme de la levure étudié par Harden était composé d'une molécule d'adénine, de deux molécules de ribose, de deux groupes phosphates et d'une molécule de *nicotinamide*. C'était en fait inhabituel de trouver cette dernière molécule dans les tissus vivants, et bien entendu l'intérêt se centra sur la nicotinamide (on l'appelle ainsi parce qu'elle contient un groupe amide $CONH_2$ et peut être obtenue facilement à partir de l'acide nicotinique. L'acide nicotinique a une structure voisine de l'alcaloïde du tabac, la *nicotine*, mais a des propriétés totalement différentes ; l'acide nicotinique est nécessaire à la vie, tandis que la nicotine est un poison mortel.) Les formules de la nicotinamide et de l'acide nicotinique sont les suivantes :



Acide nicotinique



Nicotinamide

Une fois connue la formule du coenzyme de Harden, celui-ci a été rebaptisé *diphosphopyridine nucléotide* (DPN) : *nucléotide* pour l'arrangement caractéristique d'adénine, ribose et phosphate, semblable à celui des nucléotides qui composent les acides nucléiques ; et *pyridine* à cause du nom donné à la combinaison d'atomes qui constituent le cycle d'atomes dans la formule de la nicotinamide.

Un autre coenzyme tout à fait semblable a été trouvé, différent du DPN seulement parce qu'il contient trois groupes phosphates au lieu de deux. Il a bien entendu été appelé *triphosphopyridine nucléotide* (TPN). Le DPN et le TPN sont des coenzymes pour un certain nombre d'enzymes du corps humain, qui servent tous à transférer les atomes d'hydrogène d'une molécule à une autre (ces enzymes sont appelées *déhydrogénases*). C'est le coenzyme qui fait le vrai travail de transfert de l'hydrogène ; l'enzyme elle-même choisit un substrat particulier sur lequel l'opération doit avoir lieu. L'enzyme et le coenzyme ont tous deux une fonction vitale ; et si l'un ou l'autre n'est pas fourni en quantité suffisante, la production d'énergie à partir des aliments par l'intermédiaire du transfert d'hydrogène ralentit jusqu'à devenir très faible.

Ce qui apparaît immédiatement, c'est que le groupe nicotinamide représente la seule partie de l'enzyme que le corps ne puisse fabriquer lui-même. Alors qu'il peut fabriquer toutes les protéines dont il a besoin et tous les ingrédients du DPN et du TPN sauf la nicotinamide, il doit obtenir celle-ci préfabriquée (ou au moins sous la forme d'acide nicotinique) dans le régime alimentaire. Sinon, la fabrication de DPN et de TPN s'arrête, et toutes les réactions de transfert d'hydrogène qu'ils contrôlent ralentissent.

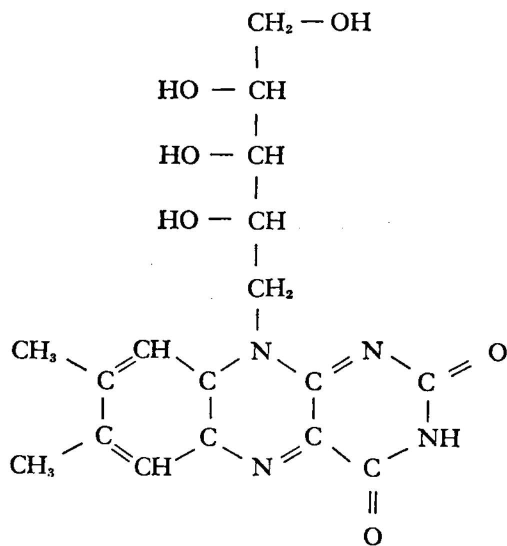
L'acide nicotinique et la nicotinamide sont-ils des vitamines ? Funk (qui inventa le mot *vitamine*) avait isolé aussi l'acide nicotinique à partir des cosses de riz. Mais l'acide nicotinique n'étant pas la substance qui guérissait le béribéri, il ne s'en est pas occupé. Toutefois, en raison de la ressemblance de l'acide nicotinique avec d'autres coenzymes, le biochimiste Conrad Arnold Elvehjem de l'université du Wisconsin l'essaya pour guérir une autre maladie par carence.

Vers 1920, le médecin américain Joseph Goldberger étudiait la *pellagre*, une maladie endémique dans les régions méditerranéennes, et presque épidémique dans le Sud des États-Unis au début du siècle. Les symptômes les plus visibles de cette maladie sont une peau sèche et râpeuse, des diarrhées et une langue enflammée ; elle conduit parfois à des troubles mentaux. Goldberger s'aperçut que cette maladie frappait des sujets dont le régime était limité (par exemple à base de maïs), et qu'elle ne touchait pas les familles possédant des vaches. Il commença donc à expérimenter des régimes artificiels sur des animaux et des détenus de prison (où la pellagre sévissait). Ayant provoqué chez des chiens une maladie semblable à la pellagre, le *blacktongue*, il réussit à guérir celle-ci grâce à l'addition d'extrait de levure. Il découvrit qu'il pouvait aussi guérir les détenus de la pellagre en ajoutant du lait dans leur régime. Goldberger décida qu'une vitamine était impliquée et l'appela le facteur P-P (antipellagre).

Par la suite, ce fut sur la pellagre que Elvehjem décida de tester l'acide nicotinique. Il en donna à manger une petite quantité à un chien atteint de « blacktongue », et le chien répondit par une remarquable amélioration. Quelques doses supplémentaires le guérirent. L'acide nicotinique était bien une vitamine, c'était le facteur P-P.

L'Association médicale américaine, craignant que le public pense que le tabac contenait des vitamines, insista pour que la vitamine ne soit pas appelée acide nicotinique, et suggéra en remplacement le nom de *niacine* (une abréviation de l'acide nicotinique) ou *niacinamide* ; la niacine est le nom courant.

Peu à peu, il apparaissait que les différentes vitamines faisaient tout simplement partie des coenzymes, et consistaient chacune en un groupe moléculaire que l'animal ou l'être humain ne pouvait fabriquer lui-même. Warburg découvrit en 1932 un coenzyme jaune catalysant le transfert des atomes d'hydrogène, et peu après le chimiste australien Richard Kuhn et ses associés isolèrent la vitamine B_2, de couleur jaune, et élucidèrent sa structure :



La chaîne de carbone attachée au cycle central ressemble à une molécule appelée *ribitol ;* on donna donc à la vitamine B_2 le nom de *riboflavine, flavine* venant du mot latin qui signifie « jaune ». Comme l'examen de son spectre montrait que la riboflavine était de la même couleur que le coenzyme jaune de Warburg, Kuhn testa le coenzyme pour l'activité riboflavine en 1935, et découvrit que celle-ci était présente. La même année, le biochimiste suédois Hugo Theorell détermina la structure du coenzyme jaune de Warburg et démontra que c'était bien de la riboflavine, mais avec un groupe supplémentaire (on a aussi montré en 1954 qu'une deuxième enzyme plus compliquée contenait une molécule de riboflavine).

Kuhn reçut le prix Nobel de chimie en 1938 et Theorell le prix Nobel de médecine et de physiologie en 1955. Toutefois, Kuhn eut le malheur d'être choisi pour ce prix juste après que l'Autriche fut envahie par l'Allemagne nazie, et il fut contraint (comme Gerhard Domagk) de le refuser.

La riboflavine a été synthétisée indépendamment par le chimiste suisse Paul Karrer. Karrer a reçu pour ce travail et d'autres travaux sur les vitamines une part du prix Nobel de chimie en 1937. (Il le partagea avec le chimiste anglais Walter Norman Haworth, qui travaillait sur la structure des cycles de carbone dans les molécules de carbohydrate.)

En 1937, les biochimistes allemands Karl Heinrich Adolf Lohmann et P. Schuster découvrirent un coenzyme important contenant de la thiamine dans sa structure. Pendant les années quarante, on observa d'autres relations entre les vitamines B et les coenzymes. La pyridoxine, l'acide pantothénique, l'acide folique, la biotine – chacune de ces vitamines était associée à un groupe d'enzyme ou à un autre.

Les vitamines illustrent bien l'économie des mécanismes chimiques du corps humain. La cellule humaine peut se dispenser de les fabriquer parce qu'elles ne servent qu'à des fonctions uniques et très spécialisées ; la cellule peut donc prendre le risque de trouver les quantités nécessaires à cela dans la nourriture. Il y a de nombreuses autres substances vitales dont le corps a besoin en quantité infime, mais qu'il doit fabriquer lui-même ; par exemple l'ATP, fabriqué à partir des mêmes éléments de base que les acides nucléiques, qui eux sont indispensables. On ne peut concevoir que l'organisme perde l'enzyme nécessaire à la synthèse des acides nucléiques tout en restant en vie, parce que les acides nucléiques sont nécessaires en quantité telle que l'organisme ne peut se fier aux aliments pour sa provision de matériaux de base indispensables. Par conséquent, aucun organisme incapable de fabriquer son propre ATP n'est connu, et ne sera jamais trouvé, selon toute probabilité.

Fabriquer des produits spéciaux comme les vitamines est un peu comme mettre en marche une machine spéciale pour fabriquer des écrous et des boulons à côté d'une ligne de montage d'automobiles. On obtient les écrous et les boulons plus facilement chez un fabricant de pièces détachées, sans qu'il y ait de perte pour l'appareil d'assemblage des automobiles ; de la même façon, l'organisme obtenant ses vitamines de la nourriture gagne en espace et en matériel.

Les vitamines illustrent un autre aspect important de la vie. Toutes les cellules vivantes ont besoin de vitamines B. Les coenzymes font donc partie des mécanismes cellulaires de toutes les cellules vivantes – végétales, animales ou bactériennes. Que la cellule obtienne sa vitamine B du régime ou qu'elle la fasse elle-même, elle en a besoin pour vivre et croître. Ce

besoin universel d'un groupe particulier de substances constitue une preuve flagrante de l'unité de la vie et de son origine (probable) dans une forme de vie unique présente dans l'Océan primitif.

LA VITAMINE A

Alors que le rôle des vitamines B est maintenant bien connu, les fonctions chimiques des autres vitamines restent difficiles à résoudre. Il n'y a que pour la vitamine A que des progrès réels ont été faits.

En 1925, les physiologistes américains L. S. Fridericia et E. Holm ont découvert que des rats nourris d'un régime déficient en vitamine A avaient de la peine à accomplir des travaux en lumière faible. L'analyse de leur rétine montra qu'il leur manquait une substance appelée *pourpre visuel.*

Il y a deux espèces de cellules dans la rétine de l'œil – les bâtonnets et les cônes. Les bâtonnets sont spécialisés dans la vision en lumière faible, et contiennent du pourpre visuel. Une déficience en celui-ci diminue donc la vision et il en résulte ce qu'on appelle la *cécité de nuit.*

En 1938, le biologiste de Harvard George Wald se mit à travailler sur la chimie de la vision. Il démontra que la lumière provoque la séparation du pourpre visuel – ou *rhodopsine* – en deux composants : la protéine *opsine,* et une fraction non protéique appelée *rétinène.* Le rétinène est tout à fait semblable en structure à la vitamine A.

A l'obscurité, le rétinène se recombine avec l'opsine pour former de la rhodopsine. Mais à la lumière, pendant sa séparation d'avec l'opsine, un faible pourcentage se décompose du fait de son instabilité. La provision de rétinène est réalimentée grâce à la vitamine A, qui peut être convertie en rétinène par l'enlèvement de deux atomes d'hydrogène par des enzymes. Ainsi, la vitamine A agit comme une réserve constante de rétinène. S'il manque de la vitamine A dans le régime, la réserve de rétinène et la quantité de pourpre visuel finissent par diminuer et il en résulte la cécité de nuit. Wald obtint une part du prix Nobel de médecine et de physiologie en 1967 pour ce travail.

La vitamine A doit avoir encore d'autres fonctions, parce qu'une déficience en vitamine A provoque la sécheresse des muqueuses et d'autres symptômes difficiles à rattacher aux troubles rétiniens de l'œil. Mais ces autres fonctions sont encore inconnues.

Il en est de même des fonctions chimiques des vitamines C, D, E et K.

Les éléments minéraux

On pourrait croire que quelque chose d'aussi impressionnant que la matière vivante serait constitué de matériaux rares. Or, si les protéines et les acides nucléiques sont des merveilles de la nature, les éléments qui entrent dans la composition du corps humain sont terriblement communs.

Quand les chimistes ont analysé les premiers composés organiques au début du XIXe siècle, on s'est rendu compte que le tissu vivant était composé principalement de carbone, d'hydrogène, d'oxygène et d'azote. Ces quatre éléments constituent à eux seuls environ 96 % du poids du corps humain,

auquel il faut ajouter le soufre, présent dans le corps en plus petite quantité. Si ces cinq éléments étaient incinérés, il ne resterait qu'un peu de cendre blanche, essentiellement des résidus osseux, et cette cendre serait constituée d'un ensemble d'éléments minéraux.

Il n'est pas étonnant de trouver dans cette cendre du sel, c'est-à-dire du chlorure de sodium ; après tout, le sel n'est pas simplement un condiment qui améliore le goût des aliments et dont on peut se passer comme par exemple le basilic, le romarin ou le thym. C'est un élément vital, et il suffit de goûter son sang pour comprendre que le sel est un composant de base du corps. Une preuve nous en est fournie par les animaux herbivores, manifestement peu difficiles quant à leur nourriture ; ils sont prêts à affronter beaucoup de dangers et de privations pour atteindre un terrain salifère qui leur permette de remédier au manque de sel de leur régime constitué de feuilles et d'herbes.

C'est au milieu du XVIIIe siècle que le chimiste suédois Johann Gottlieb Gahn a démontré la présence de calcium dans les os, et qu'un scientifique italien, Vincenzo Antonio Menghini, a observé que le sang contenait du fer. Justus von Liebig a découvert en 1847 du potassium et du magnésium dans les tissus. Vers le milieu du XIXe siècle, les constituants minéraux du corps incluaient le calcium, le phosphore, le sodium, le potassium, le chlore, le magnésium et le fer. Ceux-ci étaient impliqués dans les procédés vitaux tout autant que n'importe quel élément normalement associé aux composés organiques.

Le cas du fer est le plus simple. S'il manque dans le régime, le sang devient déficient en hémoglobine et transporte moins d'oxygène des poumons vers les cellules. Cet état est appelé *anémie par déficience en fer.* Le patient est pâle à cause du manque de pigment rouge, et fatigué parce qu'il manque d'oxygène.

Le médecin anglais Sydney Ringer a fait la découverte suivante en 1882 : le cœur d'une grenouille pouvait être maintenu en vie en dehors du corps tout en continuant à battre, lorsqu'il était conservé dans une solution (appelée encore aujourd'hui solution de Ringer) qui contient entre autres du sodium, du potassium et du calcium à peu près dans les mêmes proportions que dans le sang de la grenouille. Tous ces éléments sont donc essentiels au fonctionnement musculaire. Un excès de calcium entraîne le muscle à se bloquer en état de contraction permanente (*calcium rigor*), tandis qu'un excès de potassium l'entraîne à se décontracter en un état de relaxation permanente (*inhibition par le potassium*). De plus, le calcium est essentiel à la coagulation sanguine. Le sang ne coagule pas en son absence et aucun autre élément ne peut lui être substitué de ce point de vue.

Parmi les éléments minéraux, il en est un, le phosphore, qui joue un rôle essentiel dans l'une des fonctions les plus cruciales et diversifiées des mécanismes chimiques de la vie (voir chapitre 13).

Le calcium est un composant majeur des os qui constitue 2 % du corps humain ; le phosphore, lui, en constitue 1 %. Les autres éléments minéraux dont j'ai parlé sont présents en proportions plus faibles ; le fer ne constitue que 0,004 % du corps (ce qui fait quand même trois grammes de fer dans les tissus d'un homme adulte de taille moyenne). Mais nous ne sommes pas à la fin de la liste ; il y a d'autres éléments minéraux qui, présents dans les tissus en quantités à peine détectables, sont quand même essentiels à la vie.

La seule présence d'un élément ne signifie pas nécessairement qu'il a un rôle à jouer ; il s'agit peut-être seulement d'une impureté. Nous absorbons dans notre nourriture des quantités infimes de tous les éléments de notre environnement, dont de petites quantités se retrouvent ensuite dans nos tissus. Cependant, des éléments comme le titane et le nickel, par exemple, n'apportent rien à ces tissus, tandis que le zinc au contraire est vital. Comment distinguer entre un élément minéral essentiel et une impureté accidentelle ?

Le meilleur moyen est de démontrer que cet élément présent à l'état de trace est essentiel à une enzyme indispensable (je choisis une enzyme parce qu'il n'y a pas d'autre raison pour qu'un élément joue un rôle important à l'état de trace). En 1939, David Keilin et Thaddeus Robert Rudolph Mann, des Anglais, ont démontré que le zinc faisait partie intégrante de l'enzyme appelée anhydrase carbonique. L'anhydrase carbonique joue un rôle de premier ordre dans la distribution du gaz carbonique dans le corps, composé dont il est nécessaire de se débarrasser et dont la répartition est essentielle à la vie. La conséquence théorique est que le zinc doit être indispensable aux procédés vitaux, et c'est effectivement ce que démontre l'expérience. Par exemple, des rats nourris d'un régime pauvre en zinc s'arrêtent de grandir, perdent leurs poils, ont une peau desquamée, et meurent prématurément de ce manque de zinc, comme d'un manque de vitamine.

On a démontré de même que le cuivre, le manganèse, le cobalt et le molybdène sont indispensables à la vie animale. Leur absence dans le régime donne lieu à des maladies par carence. Le molybdène est associé à une enzyme appelée *xanthine oxydase.* L'importance du molybdène a été constatée en 1940 grâce à des travaux sur les végétaux : les spécialistes de la biologie du sol ont découvert que leurs plantes ne poussaient pas bien sur des sols dépourvus de cet élément. Il fait apparemment partie de certaines enzymes des micro-organismes du sol qui catalysent la conversion de l'azote de l'air en composés nitrifiés ; les plantes ont besoin de ces micro-organismes parce qu'elles ne peuvent pas utiliser elles-mêmes l'azote de l'air. Ceci est l'un des innombrables exemples de l'interdépendance étroite des différentes formes de vie sur notre planète. Le monde vivant n'est qu'une longue chaîne complexe sujette à bien des tribulations, voire des désastres si l'un des maillons est rompu.

Il existe des oligo-éléments qui ne sont pas indispensables. Citons le bore, qui semble nécessaire à la vie végétale, mais pas à la vie animale. Le vanadium, que certains tuniciers récoltent dans l'eau de mer et utilisent pour leur transport d'oxygène, n'est pas nécessaire à d'autres animaux. D'autres éléments comme le sélénium et le chrome sont probablement indispensables, mais leur rôle exact n'a pas encore été déterminé.

On sait à l'heure actuelle qu'il existe des déserts en oligo-éléments comme il existe des déserts dépourvus d'eau, ces deux catégories allant souvent, mais pas toujours, de pair. En Australie, les scientifiques ont découvert que trente grammes de molybdène, sous la forme d'un composé approprié répandu sur quelques hectares de sol dépourvu de molybdène, entraînent une augmentation considérable de la fertilité du sol. Le problème ne concerne pas seulement les sols exotiques. Une étude des fermes américaines en 1960 a recensé des zones déficientes en boron dans quarante et un États, en zinc dans vingt-neuf États et en molybdène dans vingt et un États. Les doses d'oligo-éléments sont déterminantes ; un excès

n'est pas préférable à une insuffisance, car des substances essentielles à la vie en petites quantités (comme le cuivre) deviennent des poisons en plus grandes quantités.

C'est poursuivre jusqu'à sa logique extrême la très ancienne coutume d'ajouter des *engrais* au sol. La pratique des engrais utilisait jusqu'à l'époque moderne des excréments animaux, guano ou fumier, qui restituent de l'azote et du phosphore au sol. Ce traitement, bien qu'efficace, s'accompagne d'odeurs désagréables et du risque inévitable de répandre des maladies infectieuses. Leur substitution par des engrais chimiques, propres et sans odeurs, a été le fruit du travail de Justus von Liebig, au début du XIXe siècle.

LE COBALT

L'un des épisodes les plus spectaculaires de la découverte des déficiences en éléments minéraux concerne le cobalt. Il se rapporte à une maladie autrefois mortelle appelée *anémie pernicieuse.*

Au début des années vingt, le pathologiste de l'université de Rochester, George Hoyt Whipple, a fait des expériences sur le renouvellement de l'hémoglobine par les aliments. Il saignait des chiens pour provoquer une anémie, et les nourrissait ensuite au moyen de divers régimes pour déterminer celui qui permettait aux chiens de remplacer l'hémoglobine perdue le plus rapidement. Il ne s'intéressait pas à l'anémie pernicieuse ni aux autres formes d'anémie, mais il travaillait sur des pigments biliaires, composés produits par le corps à partir de l'hémoglobine. Whipple découvrit que le foie était l'aliment capable de reconstituer l'hémoglobine le plus rapidement.

En 1926, deux médecins de Boston, George Richards Minot et William Parry Murphy, reprenant les conclusions de Whipple, décidèrent d'essayer le foie comme traitement de l'anémie pernicieuse chez des malades. Le traitement fut efficace. La maladie incurable était guérie lorsque les patients mangeaient du foie en proportion importante – Whipple, Minot et Murphy partagèrent le prix Nobel de physiologie et de médecine en 1934.

Les biochimistes se mirent à la recherche de la substance du foie responsable de la guérison de l'anémie. Edwin Joseph Cohn et ses collaborateurs à la Harvard Medical School préparèrent un concentré à partir du foie, cent fois plus efficace que le foie lui-même. Il fallait, toutefois, une plus grande purification pour isoler le facteur actif. Heureusement, les chimistes des Laboratoires Merck découvrirent, dans les années quarante, que le concentré de foie pouvait accélérer la croissance de certaines bactéries ; c'était donc une preuve de la grande activité de cette préparation, et les biochimistes décomposèrent le concentré en fractions qu'ils testèrent les unes après les autres. Parce que les bactéries réagissaient au facteur du foie à peu près comme elles réagissaient à la thiamine ou à la riboflavine, par exemple, les chercheurs supposèrent que le facteur recherché était une vitamine B. Ils l'appelèrent *vitamine* B_{12}.

C'est en 1948, à l'aide de cette réponse bactérienne et des techniques de chromatographie, que l'Anglais Ernest Lester Smith et Karl August Folkers chez Merck réussirent à isoler un échantillon pur de vitamine B_{12}. Cette vitamine était une substance rouge, et les deux scientifiques

pensèrent qu'elle ressemblait à la couleur de certains composés du cobalt. On savait déjà à ce moment-là qu'une déficience en cobalt provoquait une anémie grave chez le bétail et les moutons. Smith et Folkers, ayant brûlé des échantillons de vitamine B_{12} pour en analyser la cendre, constatèrent qu'elle contenait effectivement du cobalt. Ce composé est maintenant appelé *cyanocobalamine.* C'est le seul composé contenant du cobalt à avoir été trouvé jusqu'à présent dans les tissus vivants.

Les chimistes se sont vite aperçus, en décomposant et en analysant des fragments de la molécule, que la vitamine B_{12} était un composé très compliqué, pour lequel ils ont établi la formule empirique $C_{63}H_{88}O_{14}N_{14}PCo$. Sa structure globale a ensuite été déterminée aux rayons X par une chimiste anglaise, Dorothy Crowfoot Hodgkin. Le schéma de diffraction donné par les cristaux du composé lui a permis de construire une image des *densités électroniques* le long de la molécule – c'est-à-dire des régions où la probabilité de trouver un électron est élevée et de celles où elle est faible. Des lignes tracées à travers les régions d'égale probabilité dessinent une sorte d'image de la carcasse de la molécule à partir de la forme de la molécule entière.

Mais ce n'est pas aussi simple que cela. Les molécules organiques compliquées peuvent donner des schémas de diffraction aux rayons X réellement complexes, de sorte que les opérations mathématiques nécessaires pour traduire cette diffraction en densités électroniques sont extrêmement fastidieuses. Vers 1944, les ordinateurs électroniques ont facilité l'élucidation de la formule de structure de la pénicilline. Mais la vitamine B_{12} était beaucoup plus compliquée ; Hodgkin dut faire appel à un ordinateur plus moderne et se livrer à de nombreux travaux préliminaires. Hodgkin fut toutefois récompensée pour ses efforts, puisqu'elle obtint en 1964 le prix Nobel de chimie.

La molécule de vitamine B_{12}, ou cyanocobalamine, est un cycle porphyrique décentré, dans lequel manque un des ponts de carbone qui connecte les deux petits cycles pyrroles, et qui possède des chaînes latérales complexes sur le cycle pyrrole. Elle ressemble à la molécule appelée hème, relativement plus simple à une différence près : là où le groupe hème contient un atome de fer au centre du cycle porphyrique, la cyanocobalamine contient un atome de cobalt.

La cyanocobalamine injectée dans le sang des malades souffrant d'anémie pernicieuse est active en très petites quantités. Le corps peut se contenter pour cette substance d'un millième de la quantité nécessaire pour les autres vitamines B. Tous les régimes, par conséquent, contiennent suffisamment de cyanocobalamine pour nos besoins. Si ce n'est pas le cas, les bactéries des intestins en fabriquent une certaine quantité. Mais, pourquoi alors peut-on souffrir d'anémie pernicieuse ?

Des sujets atteints de cette maladie seraient simplement incapables d'absorber suffisamment de vitamine dans leur corps à travers la paroi intestinale. Leurs fèces sont en fait riches en vitamine (dont pourtant l'absence les fait mourir). Le régime à base de foie qui en fournit des quantités particulièrement élevées permet au malade d'absorber suffisamment de cyanocobalamine pour rester en vie. Mais il lui faut cent fois plus de vitamine s'il l'absorbe oralement que si elle est injectée directement dans le sang.

Quelque chose ne fonctionne pas dans le système intestinal du patient, et empêche le passage de la vitamine à travers la paroi intestinale. On

sait depuis 1929, grâce aux recherches du médecin américain William Bosworth Castle, qu'il s'agit d'un facteur du suc gastrique. Castle a appelé ce composant *facteur intrinsèque*. En 1954, les chercheurs ont découvert dans la paroi stomacale de certains animaux un produit qui facilite l'absorption de la vitamine et qui est le facteur intrinsèque de Castle. Cette substance est, semble-t-il, absente chez les patients atteints d'anémie pernicieuse. En ajoutant une petite quantité de ce facteur à la cyanocobalamine, le patient n'a plus de difficulté à absorber la vitamine par les intestins. En fin de compte, le facteur intrinsèque est une *glycoprotéine* (protéine contenant du sucre) qui se lie à une molécule de cyanocobalamine et la transporte dans les cellules de l'intestin.

L'IODE

Revenons aux oligo-éléments. Le premier élément à avoir été découvert n'était pas un métal ; c'était de l'iode, un élément dont les propriétés ressemblent à celles du chlore. Son histoire commence avec celle de la glande thyroïde.

C'est un biochimiste allemand, Eugen Baumann, qui a découvert en 1896 que la thyroïde se distinguait par son contenu en iode, par ailleurs presque totalement absent de tous les autres tissus. Un médecin qui venait de s'installer dans le Cleveland, David Marine, s'aperçut avec étonnement que les goitres étaient fréquents dans cette région. Le goitre est une maladie qui attire le regard ; la glande thyroïde peut prendre des proportions impressionnantes ; le goitre rend ses victimes soit molles et apathiques, soit nerveuses, hyperactives, avec des yeux exorbités. C'est pour le développement de techniques chirurgicales dans le traitement des thyroïdes anormales, qui soulage l'état goitreux, que le médecin suisse Emil Theodor Kocher a obtenu le prix Nobel de médecine et de physiologie en 1909. Marine se demandait si une thyroïde hypertrophiée ne résultait pas d'une déficience en iode, le seul élément dans lequel se spécialise la thyroïde, et si les goitres ne pouvaient pas être traités mieux et plus rapidement par la chimie que par le scalpel. La déficience en iode et la fréquente apparition des goitres dans la région de Cleveland allaient peut-être de pair, puisque Cleveland se trouvait à l'intérieur des terres et que l'iode, abondant dans le sol près de l'Océan et dans les fruits de mer, pouvait y manquer.

Le médecin expérimenta d'abord sur des animaux, et se sentit au bout d'une dizaine d'années suffisamment sûr de lui pour essayer de nourrir les malades atteints de goitres avec des composés contenant de l'iode. Il suggéra ensuite que des composés contenant de l'iode soient ajoutés au sel de table et à la provision d'eau des villes de l'intérieur, où le sol était pauvre en iode. Sa proposition rencontra une forte opposition, et il fallut attendre dix ans avant que l'eau et le sel soient iodés sans difficultés. Une fois que les suppléments en iode devinrent chose courante, le goitre simple en tant que maladie humaine diminua d'importance.

LES FLUORURES

Un demi-siècle plus tard, les chercheurs américains (et le public) étaient engagés dans l'étude et la discussion d'une question de santé assez

semblable : le traitement de l'eau par le fluor pour prévenir les caries dentaires. C'était un sujet à controverse dans l'arène non scientifique et politique, qui rencontra une opposition beaucoup plus farouche que dans le cas de l'iode. Peut-être la raison en était que les caries dentaires ne sont pas considérées comme aussi graves que les goitres.

Au début de ce siècle, les dentistes ont constaté que dans certaines régions des États-Unis (par exemple l'Arkansas), on avait tendance à avoir des dents noires – par marbrure de l'émail. La cause en fut finalement décelée : un contenu plus élevé que la normale des dérivés du fluor (les *fluorures*) dans l'eau potable naturelle de ces régions. Tandis que l'attention des chercheurs s'était portée sur les fluorures de l'eau, une autre découverte étonnante fut faite. Là où le contenu de l'eau en fluorures était au-dessus de la moyenne, la population avait un taux inhabituellement bas de caries dentaires. Par exemple, la ville de Galesburg en Illinois, dont l'eau contenait des fluorures, n'avait qu'un tiers du nombre de caries par enfant par rapport à la ville voisine de Quincy, où l'eau ne contenait presque pas de fluor.

La carie dentaire n'est pas une plaisanterie. Elle coûte à la population américaine plus d'un milliard et demi de dollars par an, et arrivés à l'âge de trente-cinq ans, deux tiers des Américains ont perdu au moins quelques-unes de leurs dents. Les chercheurs en odontologie ont obtenu les moyens d'étudier à grande échelle le problème du traitement de l'eau par le fluor, les risques éventuels ainsi que la prévention de la carie dentaire. Ils en ont conclu qu'une part par million de fluorure dans l'eau de consommation courante, pour un coût estimé à cinq à dix centimes par an et par personne, ne marbre pas les dents, et a cependant un effet sur la prévention des caries. Ils ont donc adopté une part par million comme standard pour tester la présence du fluorure dans l'eau courante.

Cette mesure a de l'effet tout d'abord sur les dents en cours de formation – c'est-à-dire chez les enfants. La présence de fluorure dans l'eau de boisson assure l'incorporation de petites quantités de fluorure dans la structure dentaire, qui apparemment rend le minéral des dents désagréable aux bactéries. L'usage de petites quantités de fluorure sous la forme de remèdes ou de dentifrice est également efficace pour protéger contre les caries dentaires.

La profession dentaire est maintenant convaincue, sur la base d'un quart de siècle de recherche, que pour quelques centimes par personne et par an, la carie dentaire peut être réduite d'environ deux tiers, ce qui économise au moins un milliard de dollars par an en soins dentaires et évite de la douleur et maints autres ennuis ne pouvant se chiffrer en argent.

Deux arguments principaux ont été opposés à l'utilisation des fluorures à grande échelle. Tout d'abord, les fluorures seraient toxiques, ce qui n'est pas vrai aux doses utilisées pour le traitement des eaux par le fluor. Ensuite, le traitement serait une contrainte médicale empiétant sur la liberté de l'individu, ce qui est peut-être le cas, mais ne justifie pas qu'un individu dans une société quelle qu'elle soit ait la liberté d'exposer les autres à une maladie pour laquelle il existe une prévention. Si la médication obligatoire est mauvaise, alors il faut incriminer aussi les chlorures, l'iode, et toutes les formes de vaccinations, y compris contre la variole, rendues obligatoires dans les pays civilisés de nos jours.

Les hormones

Enzymes, vitamines, oligo-éléments, à quel point ces substances peuvent décider de la vie ou de la mort de l'organisme ! Il y a un quatrième groupe de substances, en un sens encore plus puissantes, qui orchestrent l'ensemble.

Au tournant de ce siècle, deux physiologistes anglais, William Maddock Bayliss et Ernest Henry Starling, furent intrigués par un exploit du tube digestif. La glande située derrière l'estomac, appelée pancréas, relâche un suc digestif dans la partie supérieure de l'intestin exactement au moment où la nourriture quitte l'estomac pour entrer dans celui-ci. Comment le pancréas reçoit-il le message ? Qu'est-ce qui lui dit que le moment adéquat est arrivé ? A priori, on pouvait penser que l'information était transmise par le système nerveux, le seul moyen de communication alors connu dans le corps. Il était probable que l'arrivée de la nourriture dans les intestins à partir de l'estomac stimulait des terminaisons nerveuses qui faisaient passer le message au pancréas par le moyen du cerveau ou de la moelle épinière.

Pour tester cette théorie, Bayliss et Starling supprimèrent tous les nerfs du pancréas. Mais cette manœuvre échoua, car le pancréas sécrétait encore son suc au moment précis où il le fallait.

Les chercheurs, intrigués, poursuivirent leurs recherches. Ils découvrirent en 1902 un *messager chimique.* C'était une substance sécrétée par la paroi de l'intestin. Quand ils l'injectaient dans le sang d'un animal, elle stimulait la sécrétion du suc pancréatique même lorsque l'animal n'était pas en train de manger. Bayliss et Starling en conclurent que lorsque tout se passe normalement, la nourriture qui entre dans l'intestin stimule la paroi intestinale, qui sécrète une substance ; celle-ci va jusqu'au pancréas par le moyen du système circulatoire et y déclenche la sécrétion du suc pancréatique. Les deux chercheurs appelèrent la substance sécrétée par les intestins *sécrétine* et l'identifièrent comme une *hormone,* du mot grec signifiant « susciter l'activité ». On sait maintenant que la sécrétine est une petite molécule protéique.

Plusieurs années auparavant, les physiologistes avaient découvert qu'un extrait des capsules surrénales (deux petits organes situés juste au-dessus des reins) faisait augmenter la pression artérielle lorsqu'on l'injectait dans le corps. Le chimiste japonais Jokichi Takamine, qui travaillait aux États-Unis, isola la substance responsable en 1901 et l'appela *adrénaline* (ce qui devint plus tard son nom commercial, le nom chimique étant maintenant *épinéphrine*). Sa structure ressemble à l'acide aminé tyrosine, dont elle est dérivée dans le corps.

Manifestement, l'adrénaline était aussi une hormone. Avec les années, les physiologistes découvrirent qu'un certain nombre d'autres *glandes* dans le corps sécrétaient des hormones. (Le mot *glande* vient du mot grec pour « cupule » et s'applique à l'origine à n'importe quel morceau de tissu dans le corps ; mais il est devenu usuel de donner ce nom à tous les tissus qui sécrètent un liquide, même dans le cas d'organes de taille importante comme le foie ou les glandes mammaires. Les petits organes qui ne sécrètent pas de liquides ont peu à peu perdu ce nom ; c'est ainsi que les *glandes lymphatiques* ont été rebaptisées *ganglions lymphatiques.* Néanmoins, quand les ganglions lymphatiques de la gorge ou sous les bras se gonflent

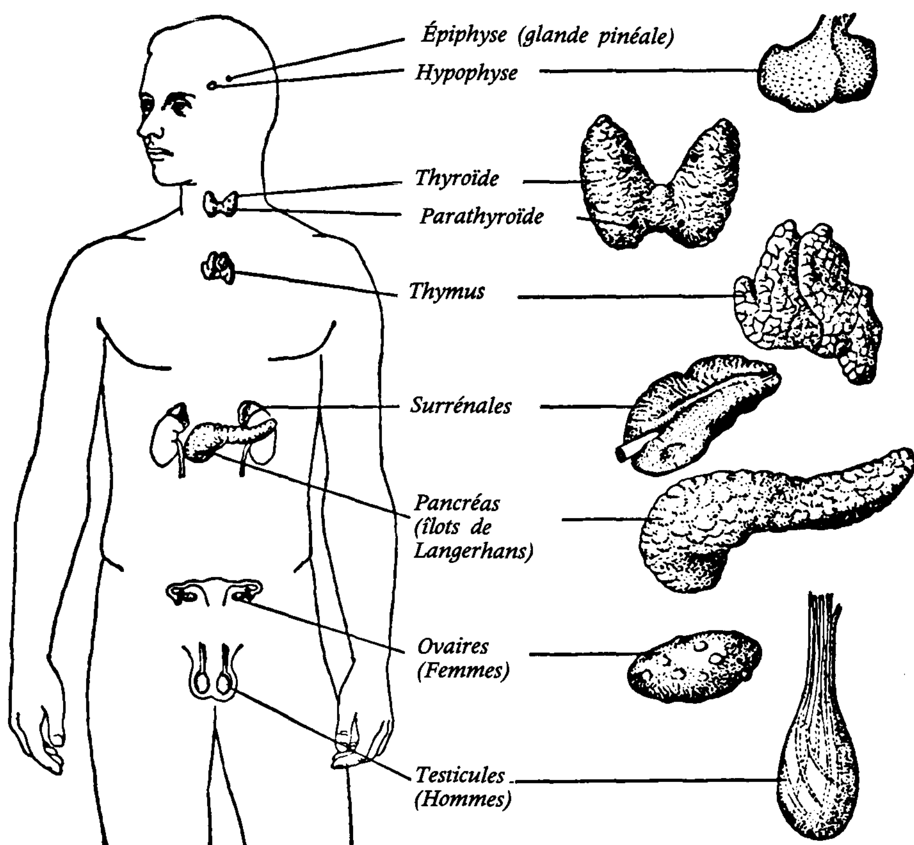
au cours d'un état infectieux, les médecins et les mères s'y réfèrent en parlant de « glandes gonflées ».)

De nombreuses glandes – comme les glandes du tractus alimentaire, les glandes sudoripares et les glandes salivaires – déchargent leur liquide par des canaux. D'autres glandes, toutefois, n'ont *pas de canaux* ; elles écoulent leurs sécrétions directement dans le sang, qui fait ensuite circuler celles-ci à travers le corps. Ce sont les sécrétions de ces glandes sans canaux, ou *endocrines*, qui contiennent les hormones (voir figure 15.1). L'étude des hormones s'appelle pour cette raison *l'endocrinologie.*

Bien sûr, les biologistes s'intéressent surtout aux hormones qui contrôlent les fonctions du corps des mammifères et des êtres humains. Toutefois, il faut mentionner le fait qu'il existe des *hormones végétales*, qui contrôlent et accélèrent la croissance des plantes, et des *hormones d'insectes*, qui contrôlent leur pigmentation, leur mue, etc.

Quand les biochimistes ont découvert que de l'iode était concentré dans la glande thyroïde, ils ont émis l'hypothèse que cet élément faisait partie d'une hormone. En 1915, Edward Calvin Kendall, de la fondation Mayo

Figure 15.1. Les glandes endocrines.



dans le Minnesota, isola de la thyroïde un acide aminé qui contenait de l'iode et qui se comportait comme une hormone ; il l'appela *thyroxine.* Chaque molécule de thyroxine contient quatre atomes d'iode. Comme l'adrénaline, la thyroxine présente une forte ressemblance de famille avec la tyrosine, et elle est fabriquée dans le corps. Plusieurs années plus tard, en 1952, la biochimiste Rosalind Pitt-Rivers et ses collègues isolèrent une autre hormone thyroïdienne – la *triiodothyronine,* ainsi appelée parce que sa molécule contient trois atomes d'iode au lieu de quatre. Elle est moins stable que la thyroxine, mais trois à cinq fois plus active.

Les hormones thyroïdiennes contrôlent le taux général de métabolisme du corps ; elles stimulent l'activité de toutes les cellules. Les gens qui ont une thyroïde sous-active sont léthargiques, mous et au bout d'un certain temps, ils deviennent des retardés mentaux parce que leurs cellules travaillent trop lentement. Inversement, ceux qui ont une thyroïde trop active sont nerveux et excités parce que leurs cellules sont suractivées. Une thyroïde trop peu active de même qu'une thyroïde trop active produisent un goitre.

La thyroïde contrôle le *taux de métabolisme de base* du corps, c'est-à-dire le taux de consommation d'oxygène au repos complet dans des conditions d'environnement confortables – le « taux au ralenti » pour ainsi dire. Si le métabolisme de base d'une personne est au-dessus ou au-dessous de la moyenne, on peut mettre en cause la glande thyroïde. La mesure du taux de métabolisme de base était autrefois pénible, parce que le patient devait jeûner avant et rester au repos une demi-heure tandis que l'on mesurait le taux, sans parler de la période qui précédait ces examens. Pour éviter tous ces inconvénients, pourquoi ne pas aller droit au but, c'est-à-dire mesurer la quantité d'hormone qui contrôle le taux produite par la thyroïde ? Ces dernières années, les chercheurs ont développé une méthode pour mesurer la *quantité d'iode associée aux protéines* (PBI) dans la circulation sanguine ; elle indique le taux de production de l'hormone thyroïdienne et fournit ainsi un test sanguin simple et facile qui remplace la mesure du métabolisme de base.

INSULINE ET DIABÈTE

L'hormone la mieux connue est l'insuline, la première protéine dont la structure ait été complètement élucidée (voir chapitre 12). Sa découverte a été le point d'arrivée d'une longue chaîne d'événements.

Le diabète est le nom d'un ensemble de maladies qui se caractérisent toutes par une soif anormale et par conséquent une production anormale d'urine. C'est le plus commun des défauts innés du métabolisme. Il y a un million et demi de diabétiques aux États-Unis, dont 80 % ont plus de quarante-cinq ans. C'est l'une des rares maladies auxquelles les femmes sont plus sujettes que les hommes puisque l'on compte quatre femmes diabétiques pour trois hommes.

Le mot « diabète » vient du grec et signifie « siphon » (il semble que celui qui inventa ce mot pensait à de l'eau siphonnant sans fin à travers le corps). La forme la plus sérieuse de la maladie est appelée diabète sucré *(diabetes mellitus). Mellitus* vient du mot grec pour « miel » et se réfère au fait qu'aux stades avancés de la maladie, dans certains cas, l'urine a un goût sucré (elle a tendance à attirer les mouches). En 1815, le chimiste français Michel

Eugène Chevreul a réussi à démontrer que cette teneur en sucre était due à la présence de glucose. Cet excès de glucose indique manifestement que le corps n'utilise pas sa nourriture de façon efficace. En fait, le patient diabétique, en dépit d'un appétit accru, perd du poids régulièrement au cours de la maladie. Il y a une génération encore, on ne connaissait pas de traitement efficace contre cette maladie.

Au XIXe siècle, les physiologistes allemands Joseph von Mering et Oscar Minkowski ont découvert que l'ablation du pancréas chez le chien produisait un état semblable au diabète humain. Après que Bayliss et Starling eurent identifié l'hormone sécrétine, il était évident qu'une hormone du pancréas était impliquée dans le diabète. Mais la seule sécrétion connue dans le pancréas était le suc gastrique. D'où venait donc cette hormone ? Un indice important fut trouvé : quand le conduit du pancréas était ligaturé, de façon qu'il ne déverse pas ses sécrétions digestives, la partie principale de la glande se ratatinait, mais le groupe de cellules connues sous le nom d'*îlots de Langerhans* (d'après le médecin allemand Paul Langerhans, qui les découvrit en 1869) restait intact.

En 1916, un médecin écossais, Albert Sharpey-Schafer, suggéra donc que ces îlots devaient produire une hormone antidiabétique. Il appela l'hormone présumée *insuline,* du mot latin pour « île ».

Les premières tentatives pour extraire l'hormone du pancréas échouèrent lamentablement. Comme nous le savons maintenant, l'insuline est une protéine, et les enzymes protéolytiques du pancréas la détruisaient avant même que les chimistes aient pu l'isoler. En 1921, le médecin canadien Frederick Grant Banting et le physiologiste Charles Herbert Best (qui travaillaient dans le laboratoire de John James Rickard MacLeod à l'université de Toronto) essayèrent une nouvelle méthode. Ils ligaturèrent le conduit du pancréas. La partie de la glande qui produit des enzymes régressa et la production d'enzymes protéolytiques s'arrêta ; les scientifiques purent alors extraire l'hormone intacte des îlots. Elle était réellement efficace contre le diabète, et on estime que dans les cinquante années suivantes, on sauva grâce à elle la vie de vingt à trente millions de diabétiques. Banting appela l'hormone *isletine,* mais la forme plus ancienne et plus latinisée proposée par Sharpey-Schafer fut finalement conservée. On l'appelle à l'heure actuelle insuline.

En 1923, Banting et Macleod (dont le principal rôle, dans la découverte de l'insuline, avait été de prêter son laboratoire pendant l'été alors qu'il était en vacances) reçurent le prix Nobel de médecine et de physiologie.

L'effet de l'insuline dans le corps se voit surtout grâce à la concentration de glucose dans le sang. D'ordinaire, le corps stocke la plupart de son glucose dans le foie, sous forme d'une sorte d'amidon appelée *glycogène* (découvert en 1856 par le physiologiste français Claude Bernard), ne laissant que de petites quantités de glucose dans la circulation sanguine pour les besoins immédiats des cellules en énergie. Si la concentration en glucose dans le sang devient trop élevée, le pancréas est stimulé à augmenter sa production d'insuline, qui se déverse dans la circulation et amène une diminution du taux de glucose. D'autre part, si le niveau de glucose tombe trop bas, sa concentration plus faible inhibe la production d'insuline, ce qui augmente le niveau de glucose, puis augmente la production d'insuline, puis diminue le taux de glucose, etc. C'est un exemple de ce qu'on appelle le *feedback.* Le thermostat qui contrôle le chauffage d'une maison fonctionne de la même manière.

Le feedback est probablement le mécanisme par lequel le corps maintient constant son environnement interne. Un autre exemple est celui de l'hormone produite par les glandes parathyroïdes, quatre petits corpuscules enfouis dans la glande thyroïde. La parathormone a été purifiée en 1960 par les biochimistes américains Lyman Creighton Craig et Howard Rasmussen après cinq ans de travail.

La molécule de parathormone est un peu plus grande que celle de l'insuline, étant constituée de 83 acides aminés et ayant un poids moléculaire de 9 500. L'action de l'hormone consiste à augmenter l'absorption de calcium dans l'intestin et à diminuer la perte de calcium à travers les reins. Chaque fois que la concentration en calcium dans le sang tombe en dessous de la normale, la sécrétion de l'hormone est stimulée. Celle-ci a pour effet de faire entrer plus de calcium qu'il n'en sort. Le niveau de calcium dans le sang s'élève ; cette augmentation inhibe la sécrétion de l'hormone. Ce jeu entre la concentration de calcium dans le sang et le flux de parathormone maintient le niveau de calcium à peu près au niveau nécessaire tout le temps. Et c'est une bonne chose, car une déviation même minime du niveau adéquat de la concentration en calcium peut provoquer la mort. Ainsi, l'ablation des parathyroïdes est fatale. Autrefois, les médecins enlevaient des morceaux de thyroïde pour soigner les goitres, ils n'avaient pas de scrupules à se débarrasser des parathyroïdes, qui étaient beaucoup plus petites et se voyaient moins ; mais la mort de leurs patients leur apprit à en avoir.

Parfois, l'effet de feedback est raffiné par l'existence de deux hormones fonctionnant dans des directions opposées. En 1961, D. Harold Copp de l'université de British Columbia (Canada) a démontré la présence d'une hormone thyroïdienne appelée *calcitonine,* qui diminue le niveau du calcium dans le sang en stimulant le dépôt de cet ion dans les os. Tandis que la parathormone tire dans une direction, la calcitonine tire dans une autre, et le feedback produit par le niveau de calcium dans le sang est d'autant mieux contrôlé (la molécule de calcitonine est constituée d'une seule chaîne polypeptidique de trente-deux acides aminés).

De même, pour la concentration de sucre dans le sang, contrôlée par l'insuline, une deuxième hormone, aussi sécrétée par les îlots de Langerhans, coopère. Ces îlots contiennent deux sortes de cellules distinctes, *alpha* et *bêta.* Les cellules bêta produisent l'insuline, tandis que les cellules alpha produisent le *glucagon.* L'existence du glucagon a été reconnue en 1923 ; il a été cristallisé en 1955. Sa molécule est constituée d'une seule chaîne de vingt-neuf acides aminés, et en 1958 sa structure a été complètement élucidée.

Le glucagon s'oppose à l'effet de l'insuline, de sorte que les deux forces hormonales fonctionnant dans des directions opposées déplacent très légèrement cet équilibre d'un côté ou de l'autre sous l'effet de la concentration dans le sang. Les sécrétions de l'hypophyse (dont je vais parler bientôt) ont aussi un effet d'opposition sur l'activité de l'insuline. Le physiologiste argentin Bernardo Alberto Houssay a obtenu une part du prix Nobel de médecine et de physiologie en 1947 pour cette découverte.

Le problème, dans le diabète, vient de ce que les îlots ont perdu la capacité de fabriquer suffisamment d'insuline. La concentration de glucose dans le sang dérive donc vers le haut. Quand le niveau s'élève à environ 50 % plus haut que la normale, il dépasse un *seuil rénal,* c'est-à-dire que le glucose se déverse dans l'urine. En un sens cette perte de glucose dans

l'urine est le moindre de deux maux parce que, si la concentration de glucose s'élevait encore, l'élévation correspondante de la viscosité du sang provoquerait une surcharge exagérée pour le cœur (le cœur est conçu pour pomper le sang, mais non pour pomper de la mélasse).

Le moyen classique de vérifier l'existence du diabète consiste à tester la présence de sucre dans l'urine. On chauffe par exemple quelques gouttes d'urine avec de la *solution de Benedict* (ainsi appelée d'après le chimiste américain Francis Gano Benedict) ; elle contient du sulfate de cuivre, qui lui donne une couleur bleu foncé. S'il n'y a pas de glucose dans l'urine, la solution reste bleue. Mais s'il y en a, le sulfate de cuivre est converti en *oxyde de cuivre*, qui est une substance rouge brique et insoluble.

Un précipité rouge au fond du tube à essai est donc le signe indiscutable de la présence de sucre dans l'urine, et permet généralement de diagnostiquer un diabète.

De nos jours, une méthode encore plus simple est à notre disposition. De petites bandes de papier d'environ cinq centimètres de long sont imprégnées avec deux enzymes, la glucose déshydrogénase et la peroxydase, ainsi qu'une substance organique appelée *orthotolidine*. La bande de couleur jaune est trempée dans un échantillon de l'urine du patient et ensuite exposée à l'air. S'il y a du glucose, il se combine avec l'oxygène de l'air grâce à l'effet catalytique de la glucose déshydrogénase. De l'eau oxygénée est formée au cours de cette réaction. La peroxydase dans le papier entraîne alors l'eau oxygénée à se combiner avec l'orthotolidine pour former un composé bleu foncé. En bref, si le papier jaune trempé dans l'urine devient bleu, on peut supposer que le diabète est en cause.

Quand le glucose commence à apparaître dans l'urine, l'état de diabète sucré est déjà assez avancé. Il vaut mieux prévenir cette maladie plus tôt en vérifiant le taux de glucose dans le sang avant qu'il passe le seuil rénal. Un *test de tolérance au glucose*, maintenant en usage, mesure la vitesse de diminution du taux de glucose dans le sang après que celui-ci a été augmenté en nourrissant la personne avec du glucose. En temps ordinaire, le pancréas répond par un flux d'insuline. Chez une personne en bonne santé, le niveau de sucre retombera à la normale en deux heures. Si le niveau reste élevé pendant trois heures ou plus, cela démontre une réponse ralentie de l'insuline et la personne peut se trouver à un stade précoce de diabète.

Il est possible que l'insuline ait quelque chose à voir avec le contrôle de l'appétit.

Nous sommes tous nés avec ce que certains physiologistes appellent un *appestat*, c'est-à-dire un système qui régule l'appétit comme un thermostat régule un four. Si l'appestat est réglé trop fort, on absorbe plus de calories qu'on ne peut en dépenser.

Au début des années quarante, le physiologiste Stephen Walter Ranson a démontré que les animaux deviennent obèses si l'on détruit une partie de *l'hypothalamus* (localisé dans la partie inférieure du cerveau), ce qui indique la localisation probable de l'appestat. Qu'est-ce qui contrôle l'opération ? Des *spasmes de faim*, peut-on penser. Un estomac vide se contracte par vagues, et l'entrée de la nourriture termine les contractions. Ces contractions pourraient être un signal pour l'appestat, mais ce n'est pas le cas : l'ablation chirurgicale de l'estomac n'a jamais interféré avec le contrôle de l'appétit.

Le physiologiste de Harvard Jean Mayer a émis une hypothèse plus subtile. Il pense que l'appestat répond au niveau de glucose dans le sang. Une fois la nourriture digérée, le niveau de glucose dans le sang descend lentement. Quand il tombe au-dessous d'un certain niveau, l'appestat est enclenché. Si, en réponse aux demandes subséquentes de l'appétit, on mange, le niveau de glucose dans le sang augmente momentanément, et l'appestat s'arrête.

LES HORMONES STÉROÏDES

Les hormones dont j'ai parlé jusqu'à présent sont toutes soit des protéines (comme l'insuline, le glucagon, la sécrétine et la parathormone), soit des acides aminés modifiés (comme la thyroxine, la triiodothyronine et l'adrénaline). Nous en arrivons maintenant à un groupe tout à fait différent, les hormones stéroïdes.

Leur histoire commence en 1927, lorsque deux physiologistes allemands, Bernhard Zondek et Selmar Aschheim, ont remarqué que des extraits d'urine de femmes enceintes, injectés à des souris ou à des rats femelles, les mettaient en état d'excitation sexuelle (cette découverte a amené au premier test de grossesse). Zondek et Aschheim avaient découvert une hormone, une *hormone sexuelle.*

Deux ans plus tard, des échantillons purs de l'hormone étaient isolés par Adolf Butenandt en Allemagne et Edward Adelberg Doisy à l'université de Saint Louis. On l'appela *œstrone,* du mot *oestrus,* terme désignant l'état d'excitation sexuelle chez la femelle. On observa bientôt que sa structure était celle d'un stéroïde, avec quatre cycles comme dans le cholestérol. Butenandt reçut pour sa participation à la découverte des hormones sexuelles le prix Nobel de chimie en 1939. Comme Domagk et Kuhn, il fut obligé de le refuser et ne put accepter cet honneur qu'en 1949, après la fin de la tyrannie nazie.

On sait maintenant que l'œstrone fait partie d'un groupe d'hormones sexuelles femelles appelées œstrogènes (« donnant lieu à l'œstrus »). En 1931, Butenandt isola la première hormone sexuelle mâle ou *androgène* (« donnant lieu à la virilité ») qu'il appela *androstérone.*

C'est la production des hormones sexuelles qui contrôle les changements apparaissant pendant l'adolescence : le développement de poils sur le visage de l'homme et le développement de la poitrine chez la femme, par exemple. Le cycle menstruel complexe de la femme dépend d'un jeu entre les différents œstrogènes.

Les hormones sexuelles femelles sont produites en grande partie dans les ovaires ; les hormones mâles dans les testicules.

Les hormones sexuelles ne sont pas les seules hormones stéroïdes. Le premier messager chimique non sexuel du type stéroïde a été découvert dans les capsules surrénales. Celles-ci, à vrai dire, sont des glandes doubles qui consistent en une glande interne appelée *médullosurrénale* (du mot latin pour « moelle ») et une glande externe appelée *corticosurrénale* (du mot latin pour « écorce »). C'est la médullosurrénale qui produit l'adrénaline. En 1929, les chercheurs ont découvert que des extraits du cortex maintenaient en vie des animaux après que leurs glandes surrénales furent enlevées – une opération mortelle à 100 %. Bien sûr, on s'est immédiatement mis à la recherche des *hormones corticoïdes.*

Des raisons médicales sous-tendaient cette recherche. La *maladie d'Addison* (décrite pour la première fois par le médecin anglais Thomas Addison) présente des symptômes qui ressemblent à ceux de l'ablation des surrénales. Visiblement, la maladie est provoquée par le défaut de production d'une hormone corticosurrénale. Peut-être des injections d'hormones corticosurrénales pouvaient soigner la maladie d'Addison, comme l'insuline agit sur le diabète.

Deux hommes se sont distingués dans ce travail : Tadeus Reichstein (qui devait plus tard réussir à synthétiser la vitamine C), et Edward Kendall (qui avait été le premier à découvrir l'hormone thyroïdienne, une vingtaine d'années plus tôt). Vers la fin des années trente, les chercheurs connaissaient plus d'une demi-douzaine de composés différents produits par le cortex surrénal. Au moins quatre d'entre eux avaient une activité hormonale. Kendall appela ces substances composés A, B, E, F, et ainsi de suite. Il s'avéra que toutes ces hormones étaient des stéroïdes.

Les surrénales sont de très petites glandes et il fallait les glandes d'innombrables animaux pour fabriquer suffisamment d'extraits de cortex. La seule solution raisonnable était d'essayer de synthétiser ces hormones.

Pendant la Deuxième Guerre mondiale, une fausse rumeur fit avancer à toute vitesse la recherche sur les hormones du cortex surrénalien. On racontait que les Allemands achetaient des glandes surrénales dans les abattoirs argentins pour fabriquer des hormones corticoïdes dans le but d'augmenter les performances de leurs pilotes pendant les vols à haute altitude. Ce n'était qu'une rumeur, mais elle eut pour effet de stimuler le gouvernement des États-Unis à donner la priorité à la recherche sur les méthodes de synthèse des hormones corticoïdes ; une priorité qui passa même avant celle accordée à la synthèse de la pénicilline ou aux anti-paludéens.

Le composé A fut synthétisé par Kendall en 1944 ; et l'année suivante, Merck commença à le produire en grande quantité. Il s'avéra malheureusement peu efficace contre la maladie d'Addison. Suite à un travail prodigieux, le biochimiste Lewis H. Sarrett de chez Merck synthétisa ensuite, par un procédé qui comportait trente-sept étapes, le composé E, appelé plus tard *cortisone.*

La synthèse du composé E ne provoqua que peu de remous dans les cercles médicaux. La guerre était finie ; les rumeurs concernant les effets magiques des corticoïdes sur les pilotes allemands se révélaient fausses ; et le composé A avait fait un fiasco. Voilà que le composé E faisait soudain son apparition.

Pendant vingt ans, le médecin Philip Showalter Hench de la clinique Mayo avait étudié le rhumatisme arthritique, une maladie douloureuse causant parfois une paralysie. Il suspectait le corps de posséder des mécanismes naturels pour faire face à cette maladie, qui disparaît souvent pendant une grossesse ou une crise de jaunisse. Il ne trouvait aucun facteur biochimique commun entre la jaunisse et la grossesse. Ayant essayé des injections de pigments biliaires (impliqués dans la jaunisse) et d'hormones sexuelles (impliquées dans la grossesse), il constata que ni les unes ni les autres ne soulageaient les patients arthritiques.

Toutefois, diverses preuves s'accumulaient en faveur des hormones corticoïdes ; en 1949, lorsque la cortisone fut disponible en quantité suffisante, Hench l'essaya ; et cela marchait ! La maladie n'était pas guérie, pas plus que le diabète par l'insuline, mais les symptômes étaient atténués,

ce qui pour un arthritique était déjà la manne tombant du ciel. Qui plus est, la cortisone s'avéra par la suite aider au traitement de la maladie d'Addison, là où le composé A avait échoué.

Kendall, Hench et Reichstein partagèrent le prix Nobel de médecine et de physiologie en 1950 pour leurs travaux sur les hormones corticoïdes.

Malheureusement, les effets des hormones corticoïdes sur les mécanismes du corps sont si complexes qu'il existe toujours des effets secondaires, parfois même graves. Les médecins répugnent à utiliser une thérapie par les hormones corticoïdes à moins que le besoin en soit impérieux. Des substances synthétiques apparentées aux hormones corticoïdes (certaines ayant un atome de fluor inséré dans leur molécule) sont utilisées pour essayer d'éviter les effets secondaires, mais rien de comparable à ce que l'on pouvait espérer n'a encore été trouvé. L'une des hormones corticoïdes les plus actives découvertes à ce jour est l'*aldostérone*, isolée en 1953 par Reichstein et ses collaborateurs.

L'HYPOPHYSE ET LA GLANDE PINÉALE

Où se trouve le contrôle de ces hormones si diverses et si puissantes ? Toutes ces hormones (y compris celles dont je n'ai pas parlé) exercent des effets plus ou moins importants sur le corps. Pourtant, elles sont si bien coordonnées entre elles que le corps fonctionne harmonieusement, sans à-coups ni ruptures de rythme. Il doit donc y avoir un chef d'orchestre quelque part qui dirige le jeu.

La réponse la plus exacte semble être l'hypophyse ; c'est une petite glande suspendue en dessous du cerveau (mais qui n'en fait pas partie). On l'appelait à l'origine la « glande pituitaire », car on supposait que sa fonction était de sécréter le « phlegme », en latin *pituita*. Cette hypothèse s'étant révélée fausse, les scientifiques l'ont rebaptisée *hypophyse* (du mot qui signifie « croître en dessous », c'est-à-dire sous le cerveau), mais on utilise encore parfois le terme de glande pituitaire.

La glande est composée de trois parties : le lobe antérieur, le lobe postérieur, et chez certains organismes, un petit pont connectant les deux lobes. Le lobe antérieur est le plus important, car il produit au moins six hormones (toutes des protéines de faible poids moléculaire), qui agissent spécifiquement sur d'autres glandes endocrines. En d'autres termes, l'hypophyse antérieure peut être considérée comme le chef d'orchestre qui maintient les autres glandes en fonctionnement sur le plan temporel et au niveau de leur coordination (il est intéressant de noter que l'hypophyse est localisée juste au centre du crâne, comme si elle était placée en un point de sécurité maximale).

L'un des messagers hypophysaires est l'*hormone de stimulation de la thyroïde* (TSH). Elle stimule la thyroïde par un effet de feedback, c'est-à-dire qu'elle entraîne la thyroïde à produire de l'hormone thyroïdienne. L'élévation de la concentration en hormone thyroïdienne dans le sang inhibe à son tour la formation de TSH par l'hypophyse ; la chute de TSH dans le sang à son tour diminue la production thyroïdienne ; ce qui stimule la production de TSH par l'hypophyse, et ainsi le cycle se maintient en équilibre.

De même, l'*hormone de stimulation corticosurrénale* (ou hormone corticotrophique : ACTH) régule le niveau des hormones corticoïdes. Si de l'ACTH

est injectée en supplément dans le corps, elle augmente le niveau de ces hormones et sert ainsi au même but que l'injection de cortisone elle-même. L'ACTH est donc utilisée pour soigner le rhumatisme arthritique.

Les recherches sur la structure de l'ACTH ont été poursuivies avec enthousiasme à cause du lien avec l'arthrite. Au début des années cinquante, son poids moléculaire a été déterminé, soit 20 000, mais elle est facilement dégradée en fragments plus petits (*corticotrophines*), qui ont une activité plus forte. La structure de l'un d'entre eux, composé d'une chaîne de trente-neuf acides aminés, est complètement connue ; même de plus petites chaînes sont efficaces.

L'ACTH peut modifier la pigmentation de la peau des animaux, ainsi que celle des êtres humains. Dans les maladies impliquant une surproduction d'ACTH, la peau devient foncée. On sait que chez les animaux inférieurs, en particulier les amphibiens, des hormones spéciales de pigmentation de la peau existent. Une hormone de ce type a été détectée en 1955 chez les êtres humains, parmi les produits hypophysaires. Elle est appelée *hormone de stimulation des mélanocytes* (les mélanocytes sont des cellules qui produisent un pigment cutané), habituellement abrégée en MSH.

La molécule de MSH a été élucidée en grande partie ; il est intéressant de noter que MSH et ACTH ont en commun une séquence de sept acides aminés ; structures et fonctions sont donc étroitement liées.

Pendant que nous discutons de pigmentation, il faut mentionner la *glande pinéale* (ou épiphyse), un corpuscule conique attaché, comme l'hypophyse, à la base du cerveau, ainsi appelée parce qu'elle a la forme d'une pomme de pin. La glande pinéale est de nature glandulaire, mais aucune hormone n'a pu y être localisée jusqu'à la fin des années cinquante. Les chercheurs travaillant sur le MSH ont utilisé 200 000 glandes pinéales de bœuf pour isoler finalement une faible quantité de substance qui, par injection, éclaircit la peau d'un têtard. L'hormone appelée *mélatonine* ne semble par contre pas avoir d'effet sur les mélanocytes humains.

La liste des hormones hypophysaires inclut encore deux hormones, l'ICSH (*hormone de stimulation des cellules interstitielles*) et le FSH (*hormone de stimulation folliculéenne*), qui contrôlent la croissance des tissus impliqués dans la reproduction. Il y a aussi l'*hormone lactogène,* qui stimule la production de lait.

L'hormone lactogène stimule d'autres activités succédant à la grossesse. De jeunes rats femelles auxquels on injecte cette hormone se mettent à fabriquer des nids même sans avoir mis de petits au monde. D'autre part, des souris, dont les hypophyses ont été enlevées peu avant la naissance de leurs petits, ne manifestent que peu d'intérêt pour leurs bébés. A la suite de ces expériences, les journaux ont baptisé l'hormone lactogène « hormone de l'amour maternel ».

Ces hormones hypophysaires, associées aux tissus sexuels, sont regroupées sous le nom de *gonadotrophines.* Une autre substance de ce type est produite par le *placenta* (l'organe qui sert à transférer les aliments du sang de la mère vers celui du fœtus et à transférer les déchets en direction opposée). L'hormone placentale est appelée *gonadotrophine chorionique* (en abrégé, HCG). De la deuxième à la quatrième semaine après le début de la grossesse, l'HCG est produite en quantité assez importante et apparaît dans l'urine. Quand on injecte des extraits de l'urine d'une femme enceinte à des souris, des grenouilles ou des lapins, on peut détecter des effets visibles. La grossesse peut être ainsi déterminée à un stade très précoce.

La plus spectaculaire des hormones de l'hypophyse antérieure est l'*hormone somatotrophe* (STH), plus communément connue sous le nom d'*hormone de croissance.* Son effet est général et stimule la croissance dans le corps entier. Un enfant qui ne peut pas produire une quantité suffisante de cette hormone devient un nain ; s'il en produit trop, un géant. Si le défaut qui aboutit à une surproduction d'hormone de croissance n'apparaît pas avant que la personne soit arrivée à maturité (c'est-à-dire quand les os sont tout à fait formés et durs), seules les extrémités (les mains, les pieds et le menton) deviennent disproportionnées – un état appelé *acromégalie,* du mot grec pour « grandes extrémités ». C'est cette hormone de croissance que Li (qui en détermina le premier la structure en 1966) synthétisa en 1970.

LE RÔLE DU CERVEAU

Les hormones agissent lentement. Elles doivent être sécrétées, transportées dans le sang vers un organe cible, et accumulées jusqu'à donner une concentration appropriée. L'action des nerfs est très rapide. A la fois un contrôle lent et un contrôle rapide sont nécessaires au corps dans diverses conditions, et l'action des deux systèmes est plus efficace que celle de l'un ou l'autre seulement. On ne pense pas que les deux systèmes soient entièrement indépendants.

L'hypophyse, qui est en quelque sorte une glande maîtresse, est proche du cerveau et en fait presque partie. La partie du cerveau à laquelle l'hypophyse est attachée par une mince tige est l'*hypothalamus,* et depuis 1920 on suspecte une sorte de connection entre les deux.

En 1945, le biochimiste anglais Geoffrey W. Harris a suggéré que les cellules de l'hypothalamus produisent des hormones qui peuvent être conduites par la circulation sanguine directement vers l'hypophyse. Ces hormones ont été décelées et appelées *releasing factors* (facteurs de déclenchement). Chacune d'entre elles provoque la production par l'hypophyse antérieure de l'une de ses hormones.

Ainsi, le système nerveux peut dans une certaine mesure contrôler le système hormonal.

Le cerveau, à vrai dire, apparaît de plus en plus non seulement comme un tableau de commande des cellules nerveuses organisées en un réseau complexe, mais aussi comme une usine chimique hautement spécialisée et probablement tout aussi complexe.

Le cerveau contient par exemple des récepteurs qui reçoivent des influx nerveux auxquels ils répondent généralement en produisant des sensations douloureuses. Les anesthésiques comme la morphine et la cocaïne se fixent aux récepteurs et font disparaître la douleur.

Certaines personnes, sous le coup d'une émotion forte, ne sentent pas la douleur, alors qu'elles la sentiraient dans d'autres conditions. Des produits chimiques naturels bloquent les récepteurs dans ces cas-là. En 1975, ces produits ont été isolés à partir de cerveaux animaux dans plusieurs laboratoires. Ce sont des peptides, de petites chaînes d'acides aminés ; les plus courts (*encéphalines*) sont constitués de cinq acides aminés seulement, tandis que les plus longs sont des *endorphines.*

Il se pourrait que le cerveau crée, de façon éphémère, de grands nombres de peptides différents dont chacun modifie l'action du cerveau d'une certaine façon – facilement produits, facilement détruits. Pour bien

comprendre le cerveau, il faudrait probablement l'étudier en détail à la fois chimiquement et électriquement.

LES PROSTAGLANDINES

Avant de quitter les hormones, il faut mentionner un groupe qui a pris de l'importance récemment, et qui ne contient ni acides aminés, ni noyau stéroïde.

En 1930, le physiologiste suédois Ulf Svante Von Euler a isolé une substance liposoluble de la glande de la prostate qui, en petites quantités, abaisse la pression sanguine et entraîne la contraction de certains muscles lisses. (Von Euler était le fils du lauréat du prix Nobel Euler-Chelpin ; il obtint lui aussi une part du prix Nobel de médecine et de physiologie en 1970 pour son travail sur la transmission nerveuse.) Von Euler appela la substance isolée *prostaglandine.*

On s'est aperçu qu'il ne s'agissait pas d'une seule substance mais de plusieurs. Au moins quatorze prostaglandines sont connues. Leur structure a été déterminée, et elles sont toutes apparentées aux acides gras non saturés. C'est peut-être à cause de la nécessité de former des prostaglandines que le corps, qui ne peut fabriquer ces acides gras, les exige du régime. Elles ont toutes un effet similaire sur la pression artérielle et la musculature lisse mais à des degrés différents, et leurs fonctions ne sont pas encore totalement élucidées.

L'ACTION DES HORMONES

Comment fonctionnent les hormones ?

On sait que les hormones n'agissent pas comme des enzymes. En tout cas, aucune hormone n'a été trouvée qui catalyse directement une réaction spécifique. On peut supposer qu'une hormone, si elle n'est pas elle-même une enzyme, agit sur une enzyme : soit qu'elle active, soit qu'elle inhibe l'activité enzymatique. L'insuline, l'hormone sur laquelle le plus de recherches ont été effectuées, semble définitivement associée à une enzyme appelée *glucokinase,* qui est essentielle pour la conversion du glucose en glycogène. Cette enzyme est inhibée par des extraits de l'hypophyse antérieure et du cortex surrénalien, et l'insuline peut annuler cette inhibition. L'insuline dans le sang sert à activer l'enzyme et accélère la conversion de glucose en glycogène. Ceci explique comment l'insuline diminue la concentration de glucose dans le sang.

Pourtant, la présence ou l'absence d'insuline affecte le métabolisme de tant de façons qu'il est difficile de déterminer comment cette seule action peut provoquer toutes les anomalies qui existent dans la chimie du corps chez un diabétique (il en est de même pour les autres hormones). Certains biochimistes ont donc cherché des effets plus globaux.

On a suggéré que l'insuline agissait comme un agent qui fait entrer le glucose dans la cellule. D'après cette théorie, un diabétique a un taux de glucose dans le sang élevé pour la simple raison que le sucre ne peut entrer dans ses cellules ; le malade ne peut donc pas l'utiliser. (Pour expliquer l'appétit insatiable du diabétique, Mayer, comme je l'ai déjà mentionné, pensait que le glucose dans le sang avait de la peine à entrer dans les cellules de l'appestat.)

Si l'insuline aide le glucose à entrer dans la cellule, elle doit agir sur la membrane cellulaire d'une façon ou d'une autre. Comment ? Les membranes cellulaires sont composées de protéines et de lipides. Nous pouvons imaginer que l'insuline, en tant que protéine, peut en quelque sorte modifier l'arrangement des chaînes latérales d'acides aminés d'une protéine de la membrane, et ouvrir ainsi les portes au glucose (et peut-être à beaucoup d'autres substances).

Si nous acceptons de nous satisfaire d'hypothèses de ce genre, nous pouvons alors supposer que les autres hormones agissent aussi sur les membranes cellulaires, chacune à sa façon, parce que chacune a son propre arrangement spécifique d'acides aminés. De même, les hormones stéroïdes comme les corps gras pourraient agir sur les molécules de lipides de la membrane, soit en ouvrant soit en fermant la porte à certaines substances. Visiblement, en aidant un matériel donné à entrer dans la cellule, ou en l'empêchant de le faire, une hormone exerce un effet considérable sur ce qui se passe dans la cellule. Elle peut fournir à une enzyme tout le substrat qu'il lui faut, priver une autre de matériel, et contrôler ainsi ce qui se produit dans la cellule. Si une seule hormone suffit à décider de l'entrée ou du rejet de plusieurs substances, nous pouvons imaginer à quel point la présence ou l'absence d'une hormone peut influencer le métabolisme, comme c'est le cas pour l'insuline.

Le tableau que je viens de dresser est attirant, mais vague. Les biochimistes préféreraient de beaucoup connaître les réactions exactes qui ont lieu sur la membrane cellulaire sous l'influence d'une hormone. Tout a commencé avec la découverte, en 1960, d'un nucléotide ressemblant à l'acide adénylique, excepté pour le groupe phosphate attaché à deux endroits différents de la molécule. Ceux qui l'ont trouvé, Wilbur Sutherland Junior et Theodore W. Rall, l'ont appelé *AMP cyclique.* Cyclique, parce que le groupe phosphate deux fois attaché forme un cycle d'atomes ; et AMP pour « adénine monophosphate », une autre appellation de l'acide adénylique. Sutherland a reçu le prix Nobel de médecine et de physiologie en 1971 pour ce travail.

Après sa découverte, l'AMP cyclique a été trouvé dans de nombreux tissus ; il a un effet prononcé sur l'activité de différentes enzymes et processus cellulaires. L'AMP cyclique est produit à partir de l'ATP présent partout, au moyen d'une enzyme appelée *adénylcyclase,* qui est localisée à la surface des cellules. Il existe probablement plusieurs enzymes de cette espèce, chacune ayant une activité en présence d'une hormone donnée. Autrement dit, l'activité de surface des hormones sert à stimuler une adénylcyclase, qui entraîne la production de l'AMP cyclique, ce qui modifie l'activité enzymatique dans la cellule et produit de nombreux changements.

Sans aucun doute, les détails sont très complexes, et des composés autres que l'AMP cyclique sont sans doute impliqués (peut-être les prostaglandines), mais ces connaissances peuvent tout de même servir de base de travail.

La mort

Grâce aux progrès de la médecine moderne dans la lutte contre les maladies infectieuses, le cancer, les problèmes alimentaires, chaque

individu a davantage de chances qu'autrefois de vivre jusqu'à la vieillesse. La moitié des gens de notre génération peut espérer atteindre l'âge de soixante-dix ans (sauf guerre nucléaire ou autre catastrophe importante).

C'est sans doute parce que peu de gens survivaient jusqu'à un âge avancé que, dans les temps anciens, on accordait autant de respect au « grand âge ». Dans *l'Iliade*, par exemple, on parle longuement du « vieux » Nestor et du « vieux » Priam. Nestor est décrit comme ayant vécu plus de trois générations humaines ; mais en un temps où la durée moyenne de vie ne dépassait guère vingt à vingt-cinq ans, il suffisait que Nestor ait soixante-dix ans pour avoir survécu à trois générations. C'est effectivement vieux, mais ce n'est rien d'extraordinaire selon les critères actuels. L'admiration portée au grand âge de Nestor au temps d'Homère a amené les mythologistes à supposer qu'il était âgé d'environ deux cents ans.

Pour prendre un autre exemple, la pièce de Shakespeare *Richard II* commence par ces mots redondants : « Le vieux Jean de Gand, Lancastre honoré par le temps ». Les contemporains de Jean, selon les chroniqueurs de l'époque, le considéraient également comme un vieil homme. Or, Jean de Gand n'a vécu que jusqu'à cinquante-neuf ans. Un exemple intéressant dans l'histoire des États-Unis est celui d'Abraham Lincoln. Que ce soit à cause de sa barbe, de son visage triste et ridé, ou des chansons de l'époque qui se référaient à lui comme au « père Abraham », beaucoup de gens l'imaginent vieux au moment de sa mort. On l'aurait souhaité ; mais il fut assassiné à l'âge de cinquante-six ans.

Ceci ne signifie pas que la vieillesse n'existait pas avant l'époque de la médecine moderne. Dans la Grèce ancienne, l'auteur dramatique Sophocle vécut jusqu'à quatre-vingt-dix ans ; l'orateur Isocrate jusqu'à quatre-vingt-dix-huit ans ; Cassiodore, dans la Rome du Ve siècle, mourut à quatre-vingt-quinze ans ; Enrico Dandolo, le doge de Venise, au XIIe siècle, vécut jusqu'à quatre-vingt-dix-sept ans ; le peintre de la Renaissance Le Titien vécut jusqu'à quatre-vingt-dix-neuf ans ; le duc de Richelieu, petit-neveu du célèbre cardinal, à l'époque de Louis XV, vécut jusqu'à quatre-vingt-douze ans ; et l'écrivain français Bernard Le Bovier de Fontenelle mourut tout juste avant de devenir centenaire.

Ainsi, bien que l'espérance moyenne de vie dans les sociétés médicalement avancées ait beaucoup augmenté, l'âge maximum atteint n'a pas changé. Même de nos jours, très peu d'individus atteignent ou dépassent l'âge d'un Isocrate ou d'un Fontenelle. Et on n'attend pas non plus des nonagénaires modernes qu'ils soient capables de participer à la vie active avec beaucoup de vigueur. Sophocle écrivait encore des pièces célèbres à quatre-vingt-dix ans ; Isocrate composait de grands discours ; Le Titien peignit jusqu'à la dernière année de sa vie ; et Dandolo fut le chef incontesté de la guerre vénitienne contre l'Empire byzantin à l'âge de quatre-vingt-seize ans. Parmi les vieillards relativement vigoureux des temps modernes, le meilleur exemple auquel je puisse penser est George Bernard Shaw, qui vécut jusqu'à quatre-vingt-quatorze ans, et le mathématicien anglais et philosophe Bertrand Russell, qui était encore actif lorsqu'il mourut à 98 ans.

Bien qu'une proportion plus élevée qu'autrefois de notre population atteigne l'âge de soixante ans, l'espérance de vie au-delà de cet âge s'est peu améliorée par rapport au passé. La Compagnie métropolitaine d'assurances sur la vie estimait en 1931 que l'espérance de vie d'un Américain de sexe mâle de soixante ans était environ la même qu'un siècle

et demi plus tôt – c'est-à-dire 14,3 ans au lieu de 14,8 ans autrefois. Pour la femme américaine, les chiffres correspondants sont 15,8 et 16,1. L'apparition des antibiotiques depuis 1931 a fait augmenter l'espérance de vie après soixante ans pour les deux sexes d'environ deux ans et demi. Mais dans l'ensemble, en dépit de tout ce que la médecine et la science ont pu faire, la vieillesse a raison d'une personne à peu près à la même vitesse et de la même façon que toujours. Nous n'avons pas encore trouvé le moyen de prévenir l'affaiblissement progressif et l'arrêt final de la machine humaine.

L'ARTÉRIOSCLÉROSE

Pour le corps comme pour beaucoup d'autres machines, ce sont d'abord les parties mobiles qui s'abîment. Le système circulatoire – le cœur avec ses battements et les artères – sont le talon d'Achille de l'être humain à longue échéance. Les troubles de ce système sont devenus le numéro un des causes de décès. Les maladies circulatoires sont responsables de plus de la moitié des morts aux États-Unis, et parmi ces maladies, l'une d'entre elles, *l'artériosclérose,* est responsable d'une mort sur quatre.

L'artériosclérose (du grec pour « durcissement ») se caractérise par le dépôt de petits grains de lipides le long de la surface interne des artères, qui contraint le cœur à travailler plus pour conduire le sang à travers les vaisseaux à une vitesse normale. La pression sanguine augmente et l'effort subséquent des petits vaisseaux sanguins peut les faire éclater. Si cela a lieu dans le cerveau (une zone particulièrement vulnérable), il y a hémorragie cérébrale ou *congestion cérébrale.* Parfois l'éclatement d'un vaisseau est si insignifiant qu'il ne cause qu'un malaise temporaire, ou même passe inaperçu, mais l'éclatement massif des vaisseaux amène la paralysie ou une mort rapide.

Les *artères coronaires* qui amènent l'oxygène au cœur sont particulièrement sensibles au rétrécissement par l'artériosclérose. Le manque d'oxygène qui en résulte pour le cœur donne lieu à la douleur atroce de l'*angine de poitrine,* qui aboutit finalement plus ou moins rapidement à la mort.

Le durcissement et le rétrécissement des artères introduit un nouveau danger. La friction plus importante du sang contre les surfaces internes durcies des vaisseaux augmente les chances de formation des caillots, et l'étroitesse des vaisseaux celles qu'un caillot bloque complètement le flux sanguin. Dans l'artère coronaire, qui nourrit le muscle cardiaque lui-même, un caillot peut provoquer une mort presque instantanée *(thrombose coronaire).*

La cause exacte de la formation de dépôts dans la paroi artérielle est l'objet de débats parmi les médecins. Le cholestérol pourrait en être une, mais on est loin de savoir de quelle manière. Le plasma humain contient des *lipoprotéines,* dont le cholestérol, ainsi que d'autres corps gras associés à des protéines. Certaines des fractions lipoprotéiques se maintiennent à une concentration constante dans le sang – que l'individu soit sain ou malade, avant ou après les repas, etc. D'autres sont fluctuantes et s'élèvent par exemple après les repas. D'autres encore sont beaucoup plus élevées chez les individus obèses. La fraction cholestérol est particulièrement élevée chez les gens trop gros et chez ceux qui souffrent d'artériosclérose.

L'artériosclérose, ainsi que l'obésité, a tendance à être associée à un contenu élevé du sang en lipides. Les gens trop forts sont plus sujets à l'artériosclérose que les gens minces. Les diabétiques ont eux aussi des niveaux élevés de lipides dans le sang et sont plus souvent atteints par l'artériosclérose que les individus normaux. Pour terminer le tableau, le diabète est beaucoup plus fréquent chez les gens trop gros que chez les gens minces.

Ce n'est donc pas par accident que ceux qui vivent vieux sont souvent des gens petits et maigres. Les hommes grands et forts, souvent joviaux, ne font généralement pas attendre inutilement le fossoyeur. Il y a bien sûr toujours des exceptions, par exemple Winston Churchill et Herbert Hoover, qui dépassèrent leur quatre-vingt-dixième anniversaire, bien qu'ils ne fussent pas réputés pour leur minceur.

La question la plus importante reste de savoir si l'artériosclérose est provoquée ou peut être prévenue par le régime alimentaire. Les graisses alimentaires animales – celles du lait, du beurre et des œufs – sont particulièrement riches en cholestérol, tandis que les graisses végétales n'en contiennent pas. De plus, les acides gras des graisses végétales sont de type non saturé et s'opposeraient, semble-t-il, au dépôt de cholestérol. Les recherches réalisées à ce sujet en 1984 tendent à démontrer que le cholestérol dans le régime favorise le développement de l'artériosclérose, ce qui fait que les gens suivent désormais *des régimes bas en cholestérol* dans l'espoir d'éviter l'artériosclérose.

Bien sûr, le cholestérol dans le sang n'est pas obligatoirement dérivé du cholestérol dans le régime. Le corps peut et doit fabriquer son propre cholestérol facilement ; et en se nourrissant d'un régime complètement dépourvu de cholestérol, on en a tout de même une provision importante dans les liproprotéines du sang. On peut donc supposer que ce n'est pas seulement la présence du cholestérol qui intervient, mais aussi la tendance de l'individu à le déposer là où il peut faire le plus de mal. Il pourrait y avoir des tendances héréditaires à fabriquer des excès de cholestérol. Les biochimistes recherchent des remèdes inhibant sa formation dans l'espoir de prévenir le développement de l'artériosclérose chez les sujets enclins à cette maladie.

Entre-temps, la technique chirurgicale de l'*anastomose* (on dit aussi *shunt*) est devenue très courante depuis son introduction en 1969, et elle a connu de nombreux succès : on utilise des vaisseaux du corps du patient que l'on relie aux artères coronaires en faisant circuler le sang par l'anastomose, court-circuitant ainsi la région sclérosée, de façon à assurer au cœur l'apport sanguin nécessaire. Si cette technique n'augmente pas l'espérance de vie globale, elle évite les douleurs de l'angine de poitrine à la fin d'une vie, et c'est déjà beaucoup pour ceux qui en souffrent.

LA VIEILLESSE

Nous n'échappons pas à la vieillesse même si nous échappons à l'artériosclérose. La vieillesse est une maladie universelle. Rien n'arrête ses effets progressifs vers le déclin, la friabilité croissante des os, l'affaiblissement des muscles, le durcissement des articulations, le ralentissement des réflexes, la diminution de la vue et de la vivacité du cerveau. La vitesse à laquelle tout cela se produit est un peu plus faible

chez certains que chez d'autres, mais de toute façon, le processus est inexorable.

Peut-être ne devons-nous pas trop nous en plaindre. Si la vieillesse et la mort finissent par venir, c'est avec une lenteur surprenante. En général, la durée de vie des mammifères est fonction de leur taille. Le plus petit des mammifères, la musaraigne, a une vie d'un an et demi, et le rat, de quatre ou cinq ans. Un lapin vit quinze ans, un chien dix-huit ans, un porc vingt ans, un cheval quarante ans, et un éléphant soixante-dix ans. Par ailleurs, plus l'animal est petit, plus son rythme de vie est rapide, par exemple ses battements cardiaques. Le cœur d'une musaraigne bat à mille coups par minute, tandis que celui d'un éléphant bat à vingt coups par minute.

En fait, la durée de vie des mammifères correspond à peu près à un milliard de battements cardiaques. Les êtres humains font exception à cette règle de façon frappante. Plus petit que le cheval et plus encore que l'éléphant, l'être humain vit pourtant plus longtemps qu'aucun autre mammifère. Même sans prendre en considération les récits d'hommes ayant atteint des âges incroyables – pour lesquels il n'existe pas de source précise –, nous possédons des données tout à fait valables selon lesquelles la durée de vie peut atteindre cent quinze ans. Les seuls vertébrés à battre ce record sont de grandes tortues lentes.

Les battements du cœur humain, environ soixante-douze par minute, correspondent bien à ceux d'un mammifère de sa taille. En soixante-dix ans, l'espérance moyenne de vie dans les régions du monde technologiquement avancées, le cœur humain bat deux milliards et demi de fois ; à cent quinze ans, quatre milliards de fois. Même nos plus proches parents, les grands singes, ne peuvent égaler ce record, ni s'en approcher. Le gorille, considérablement plus grand que l'homme, ne dépasse pas cinquante ans.

Il n'y a pas de doute que le cœur humain réalise un exploit supérieur au cœur des autres espèces (celui de la tortue dure peut-être plus longtemps, mais l'intensité de sa vie n'est pas comparable). On se demande pourquoi nous avons d'aussi longues années à vivre ; mais bien sûr, en tant qu'êtres humains, nous pensons surtout à nous demander pourquoi nous ne vivons pas plus longtemps.

De toute façon, qu'est-ce que la vieillesse ? Jusqu'à présent, on ne peut que spéculer. Certains ont suggéré que la résistance du corps aux infections diminue lentement avec l'âge, à une vitesse qui dépend de caractéristiques héréditaires. D'autres ont spéculé que des *résidus* de quelque sorte s'accumulent dans les cellules (là aussi à une vitesse variable selon les individus). Cette hypothèse suppose que des produits secondaires sont fabriqués dans les réactions cellulaires normales, produits que la cellule ne peut pas détruire et dont elle ne peut pas se débarrasser ; ces produits s'accumulent dans la cellule avec les années, jusqu'à ce que finalement ils interfèrent suffisamment avec le métabolisme cellulaire pour qu'il cesse de fonctionner. Selon cette théorie, quand beaucoup de cellules sont hors de fonction, le corps meurt. Une variante de cette théorie prétend que les molécules des protéines deviennent des résidus, du fait de liaisons qui se développent entre elles et les rendent rigides et cassantes, entraînant finalement l'arrêt des mécanismes cellulaires.

S'il en est ainsi, alors l'« échec » fait partie intégrante des mécanismes cellulaires. La capacité démontrée par Carrel de conserver un morceau de tissu embryonnaire en vie pendant des dizaines d'années semblerait

prouver que les cellules elles-mêmes sont peut-être immortelles, et que c'est seulement leur organisation en millions de millions de cellules individuelles qui amènerait la mort. C'est l'organisation qui échoue et non les cellules.

Apparemment, ce n'est pas le cas. On pense maintenant que Carrel a peut-être involontairement introduit des cellules neuves dans sa préparation en cherchant à nourrir les tissus. Les essais destinés à travailler sur des cellules isolées ou des groupes de cellules, qui excluent rigoureusement l'introduction de nouvelles cellules, semblent montrer que les cellules vieillissent inexorablement et ne peuvent se diviser plus de cinquante fois au total, probablement tout en subissant des changements irréversibles de certains composants clés.

Cela n'explique pas la longue durée de la vie humaine. Se peut-il que le tissu humain ait développé des méthodes pour renverser ou inhiber les effets du vieillissement cellulaire, des méthodes plus efficaces que chez les autres mammifères ? Encore une fois, les oiseaux ont tendance à vivre plus longtemps que les mammifères de la même taille, en dépit du fait que le métabolisme aviaire soit plus rapide que celui des mammifères – là encore, capacité supérieure à éviter ou inhiber la vieillesse.

Si le vieillissement peut être mieux évité par certains organismes que par d'autres, pourquoi les êtres humains n'apprendraient-ils pas la méthode pour y apporter des améliorations ? Peut-être la vieillesse est-elle guérissable, peut-être pouvons-nous développer la capacité de vivre plus longtemps – et, qui sait, devenir immortels ?

Certains sont d'un optimisme sans faille à cet égard. Les miracles médicaux du passé semblent annoncer des miracles sans fin pour le futur. Si c'est le cas, quel dommage de faire partie d'une génération qui risque de manquer de peu la guérison du cancer, de l'arthrite et du vieillissement !

Vers la fin des années soixante, un mouvement s'est créé en faveur de la congélation du corps humain au moment de la mort, de façon à conserver les mécanismes cellulaires aussi intacts que possible jusqu'aux jours bénis où la maladie à l'origine de la mort de l'individu congelé pourrait être guérie. L'individu ressusciterait et redeviendrait sain, jeune et heureux.

Bien sûr, rien, à l'heure actuelle, ne laisse penser qu'un corps mort puisse être rendu à la vie, ou qu'un corps congelé – même s'il est en vie au moment de la congélation – puisse être rendu à la vie par décongélation. Les partisans de ce procédé (la *cryonique*) ne se soucient guère des complications que pourrait créer le flot de cadavres revenant à la vie.

En fait, cela n'a pas de sens de congeler un corps intact, même en admettant qu'un retour à la vie soit possible. C'est du gâchis. Les biologistes ont eu jusqu'à présent beaucoup plus de succès en développant des organismes entiers à partir de groupes de cellules spécialisées. Les cellules de la peau ou du foie sont, après tout, équipées génétiquement comme les autres cellules et comme l'œuf fertilisé d'origine. Elles sont spécialisées parce que certains gènes sont inhibés ou activés à des degrés divers. Mais peut-être ces gènes pourraient-ils être désinhibés ou désactivés, et rendre la cellule identique à un œuf fertilisé d'où se développerait de nouveau un organisme entier – le même organisme au niveau génétique que celui dont elle avait fait partie ? Ce procédé (appelé *clonage*) offre plus d'espoir de conserver le schéma génétique (sinon la mémoire et la personnalité). Au lieu de congeler le corps entier, couper le petit orteil et le congeler.

Mais désirons-nous vraiment être immortels – soit par l'intermédiaire du congélateur, soit par clonage, soit par simple réversion du phénomène de vieillissement chez l'individu ? Il y a peu d'êtres humains qui n'accepteraient pas joyeusement l'immortalité si elle ne s'accompagnait pas de douleurs, de maladies et de la vieillesse – mais supposons que nous soyons tous immortels ?

S'il n'y avait que peu ou pas de décès sur la Terre, il n'y aurait pas non plus de naissances. Ce serait une société sans enfants. Ce qui n'est pas impossible mais, à dire vrai, une société suffisamment égocentrique pour désirer l'immortalité ne s'arrêterait pas à l'élimination des enfants.

Et ensuite ? Cette société se composerait toujours des mêmes cerveaux, qui auraient les mêmes idées, elle tournerait en rond sans fin. Souvenons-nous que les enfants n'ont pas seulement de tout jeunes cerveaux, mais aussi des cerveaux *neufs*. Chaque enfant (en dehors des cas de naissances multiples identiques) a un équipement génétique différent de celui de tous les autres êtres humains ayant jamais vécu. Grâce aux enfants, des combinaisons génétiques nouvelles sont constamment fournies à l'humanité, de sorte que la route est ouverte à l'amélioration et au développement.

S'il est avisé de diminuer le taux de naissances, devons-nous le réduire à néant ? Ce serait une bonne chose d'éliminer les douleurs et les désagréments liés à la vieillesse, mais devons-nous créer une espèce constituée de gens vieux, fatigués, moroses, toujours les mêmes, et ne jamais ouvrir la voie au neuf et au meilleur ?

Peut-être la perspective de l'immortalité est-elle pire que celle de la mort.

IV. Aspects de l'évolution



Planche IV.1. Complexe ADN-protéine photographié au microscope électronique. Les granulations, isolées à partir de cellules germinales d'un animal marin et grossies 77 500 fois, résultent d'un mélange d'ADN et de protéine. (Avec l'aimable autorisation de United Press International.)

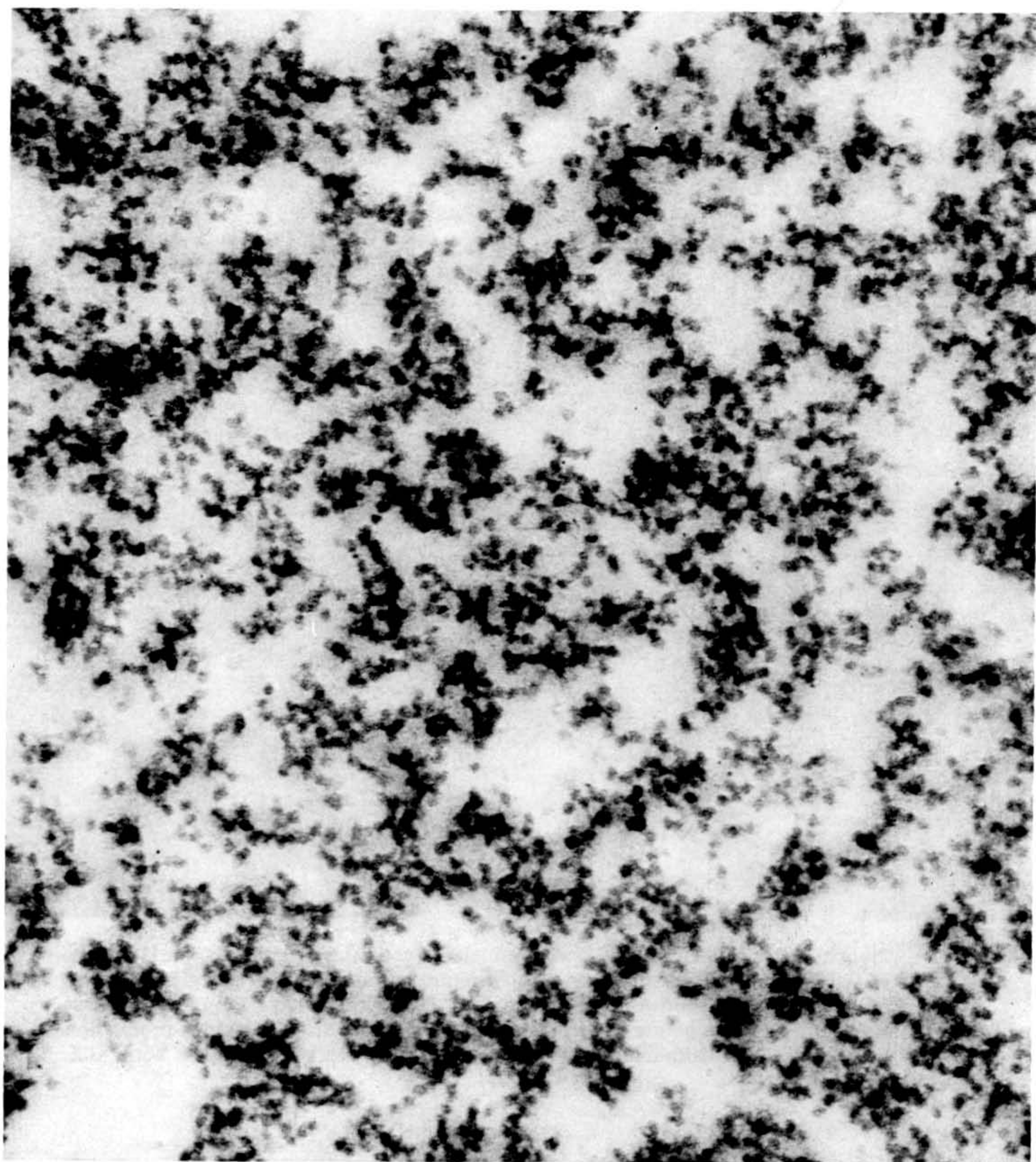


Planche IV.2. Particules de ribonucléoprotéine (ribosomes) de cellules de foie de cochon d'Inde. Ces particules sont le lieu principal de la synthèse des protéines dans la cellule. (Avec l'aimable autorisation de J.F. Kirsch, université Rockefeller, New York, 1961.)

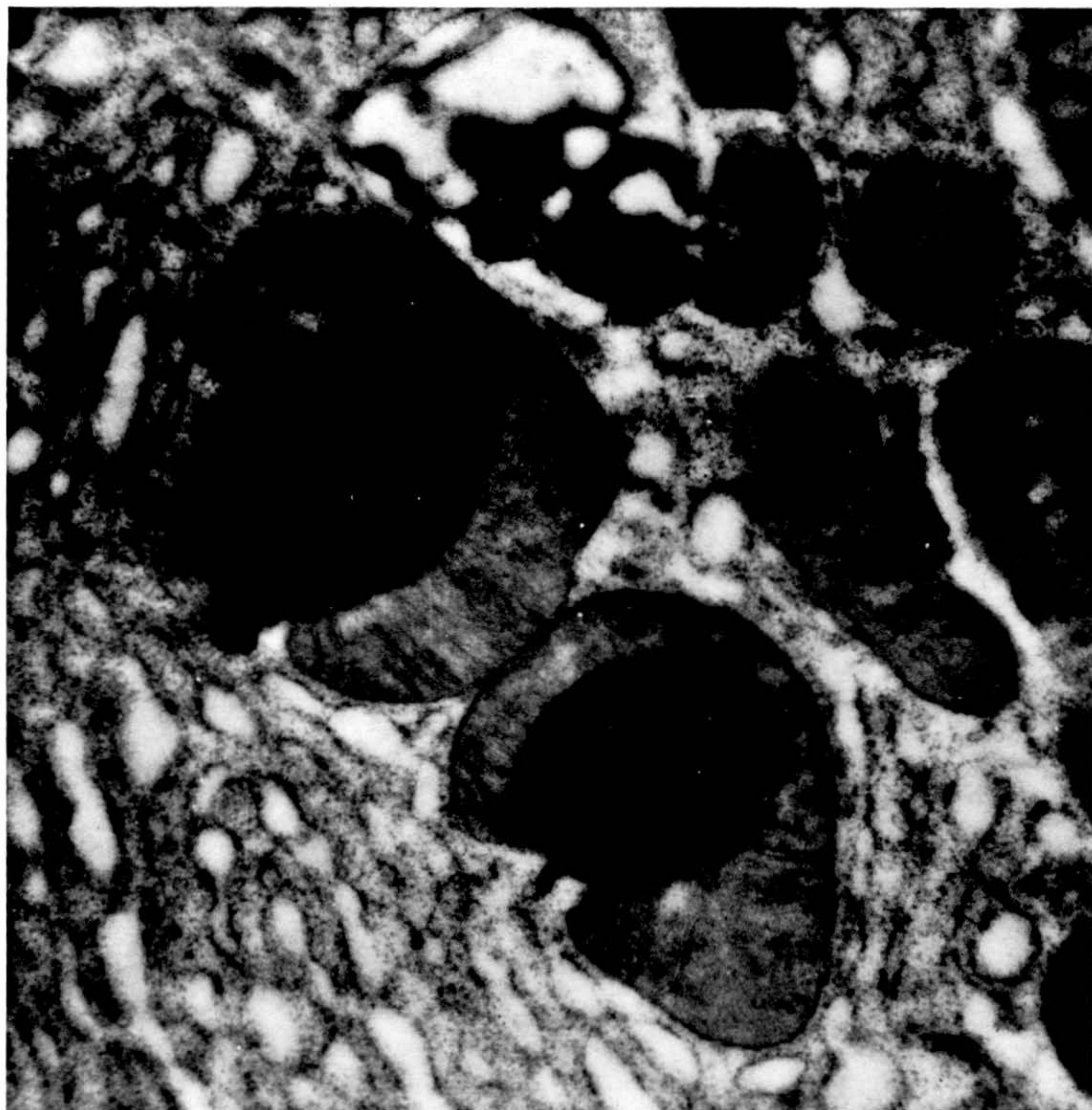


Planche IV.3. Mitochondries, parfois appelées « centrales thermiques » de la cellule, car elles sont à l'origine des réactions chimiques qui produisent l'énergie nécessaire à la vie. Les mitochondries apparaissent comme des croissants gris autour des corps noirs, qui sont des gouttes de lipides servant de « carburant » pour la production d'énergie. (Avec l'aimable autorisation de l'institut Rockefeller, New York, G.E. Palade.)

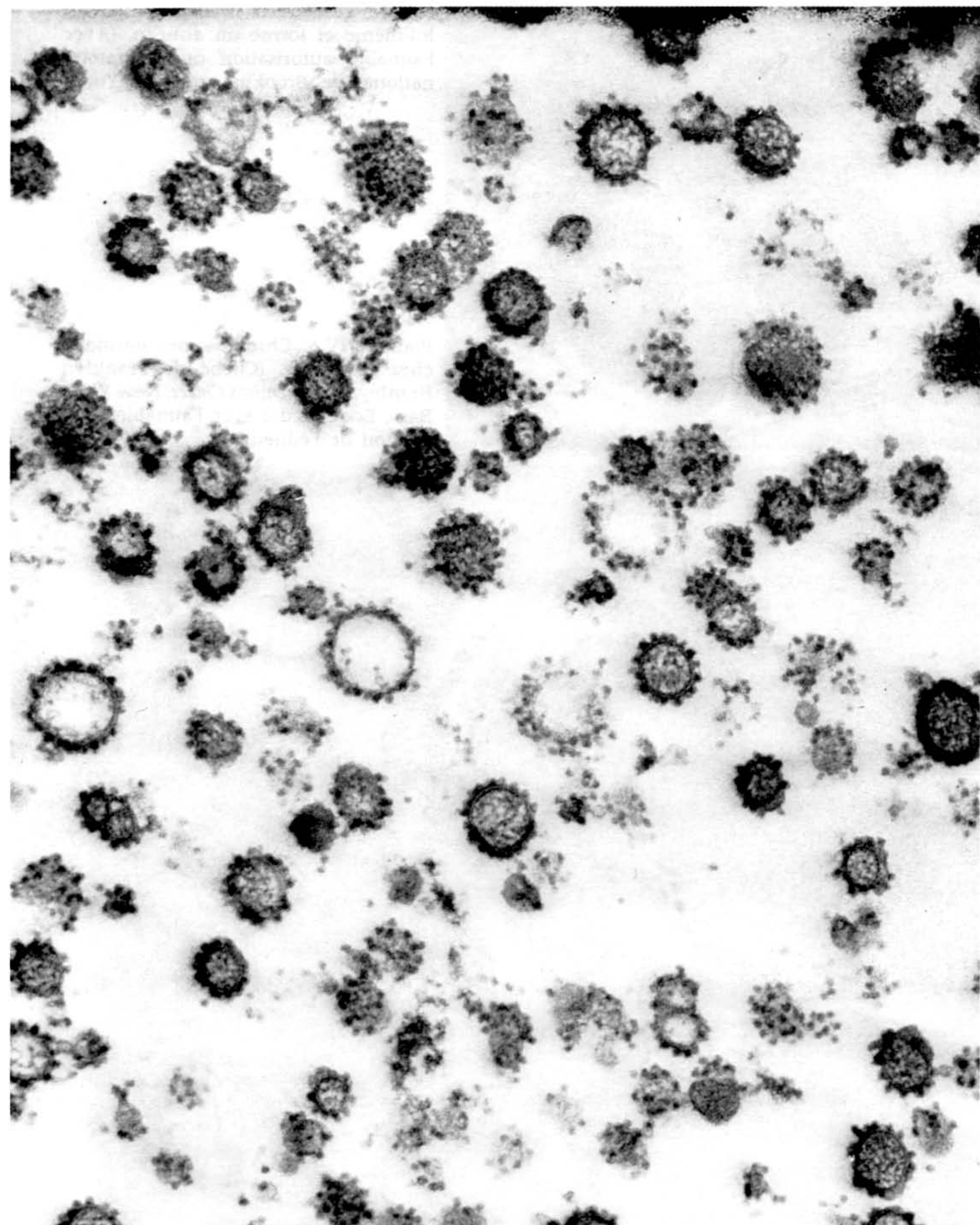


Planche IV.4. Ribosomes, corpuscules du cytoplasme de la cellule. Ceux-ci ont été obtenus par centrifugation de cellules de pancréas, et grossis environ 100 000 fois au microscope électronique. On les trouve soit libres, soit attachés à des vésicules de la membrane cellulaire, les microsomes. (Avec l'aimable autorisation de l'institut Rockefeller, New York, G.E. Palade.)

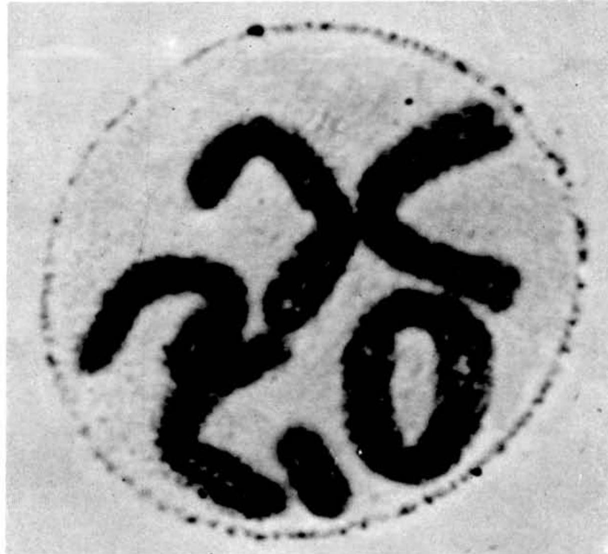


Planche IV.5. Chromosomes endommagés par des radiations. Certains sont brisés ; l'un d'eux s'est enroulé sur lui-même et forme un anneau. (Avec l'aimable autorisation du laboratoire national de Brookhaven, New York.)

Planche IV.6. Chromosomes normaux chez *Drosophila*. (Cliché de Franklyn Branley, éd., *Scientist's Choice*, New York, Basic Books, s.d. ; avec l'aimable autorisation de l'éditeur.)

Planche IV.7. Mutations chez la drosophile, ici sous la forme d'ailes recroquevillées. Les mutations ont été obtenues en exposant le parent mâle à des radiations. (Avec l'aimable autorisation du laboratoire national de Brookhaven, New York.)

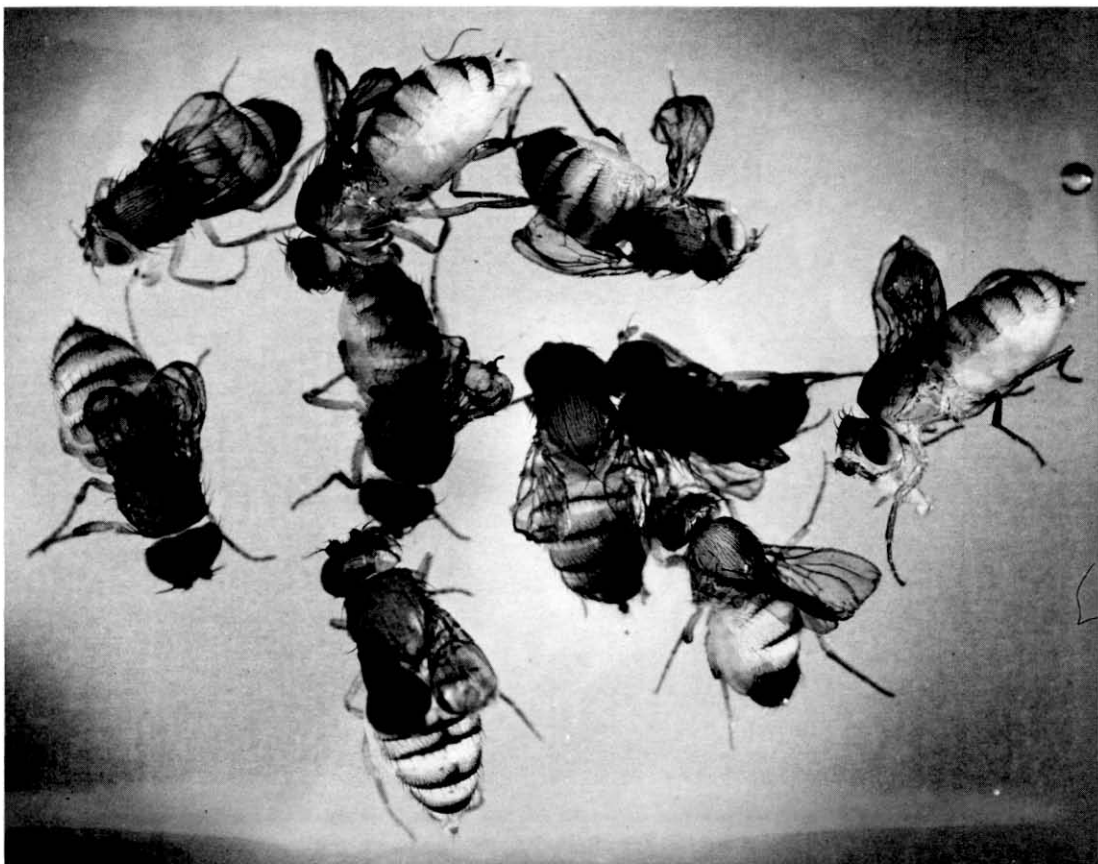




Planche IV.8. Pour étudier les effets des radiations, on a exposé de jeunes plants de rose (variété Better Times) à 5 000 roentgens de rayons gamma pendant vingt-quatre heures. A la floraison, quelques mois plus tard, on a observé un certain nombre de mutations : la fleur au centre est un mutant instable pour la couleur, alors que celle de droite est un mutant stable. A gauche, une rose témoin (non irradiée) de la même variété. (Avec l'aimable autorisation du laboratoire national de Brookhaven, New York.)



Planche IV.9. Mutation spontanée chez un chrysanthème (variété Masterpiece, rose), qui a développé un secteur de couleur bronze. Ce mutant a été baptisé Bronze Masterpiece. On a pu obtenir ce phénomène par irradiation de plants de Masterpiece, à un taux de fréquence beaucoup plus élevé que dans la nature. L'irradiation de plants de Bronze Masterpiece peut provoquer une réversion vers la couleur rose originale de Masterpiece. (Avec l'aimable autorisation du laboratoire national de Brookhaven, New York.)

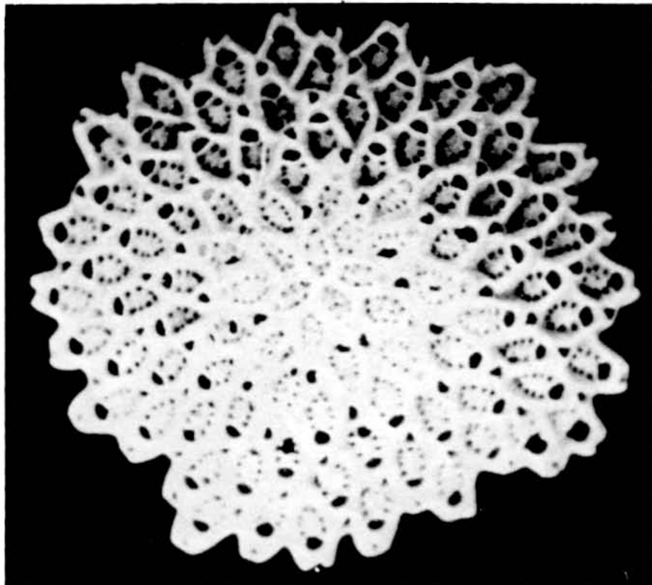


Planche IV.10. Fossile de bryozoen, petit animal aquatique qui s'apparente aux mousses, grossi environ vingt fois. Ce spécimen a été trouvé lors d'un forage au large du cap Hatteras. (Avec l'aimable autorisation de United Press International.)

Planche IV.11. Fossile de foraminifère, trouvé dans les mêmes conditions que ci-dessus. La craie et certains calcaires sont formés en grande partie de coquilles provenant de ces animaux microscopiques unicellulaires. Les falaises de Douvre et les blocs de pierre qui ont servi à la construction des pyramides d'Egypte fournissent des exemples célèbres. (Avec l'aimable autorisation de United Press International.)

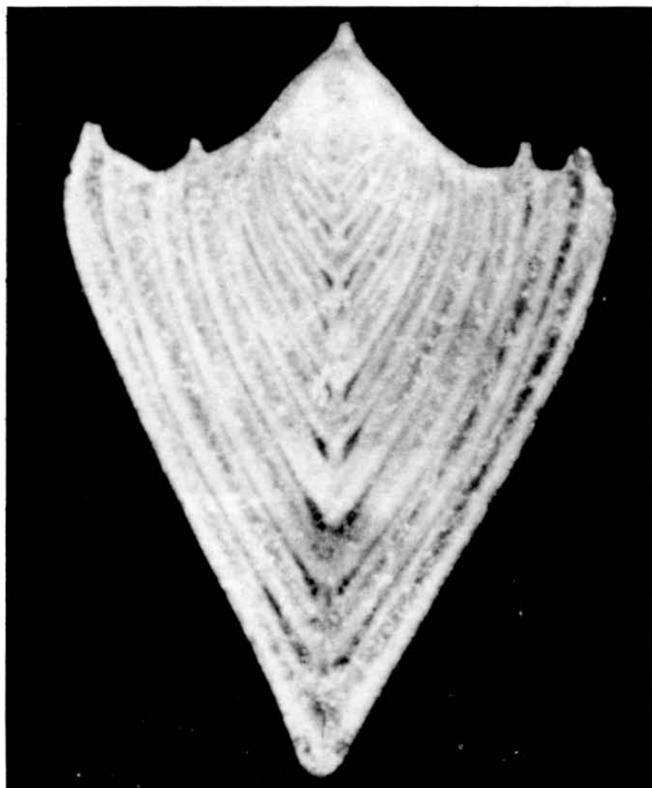




Planche IV.12. Fossile de crinoïde, ou lys des mers, animal primitif appartenant au superphylum des échinodermes. Ce spécimen provient de l'Indiana. (Nég. n° 120809, photo Thane Bierwert ; avec l'aimable autorisation de la bibliothèque du Musée américain d'histoire naturelle.)

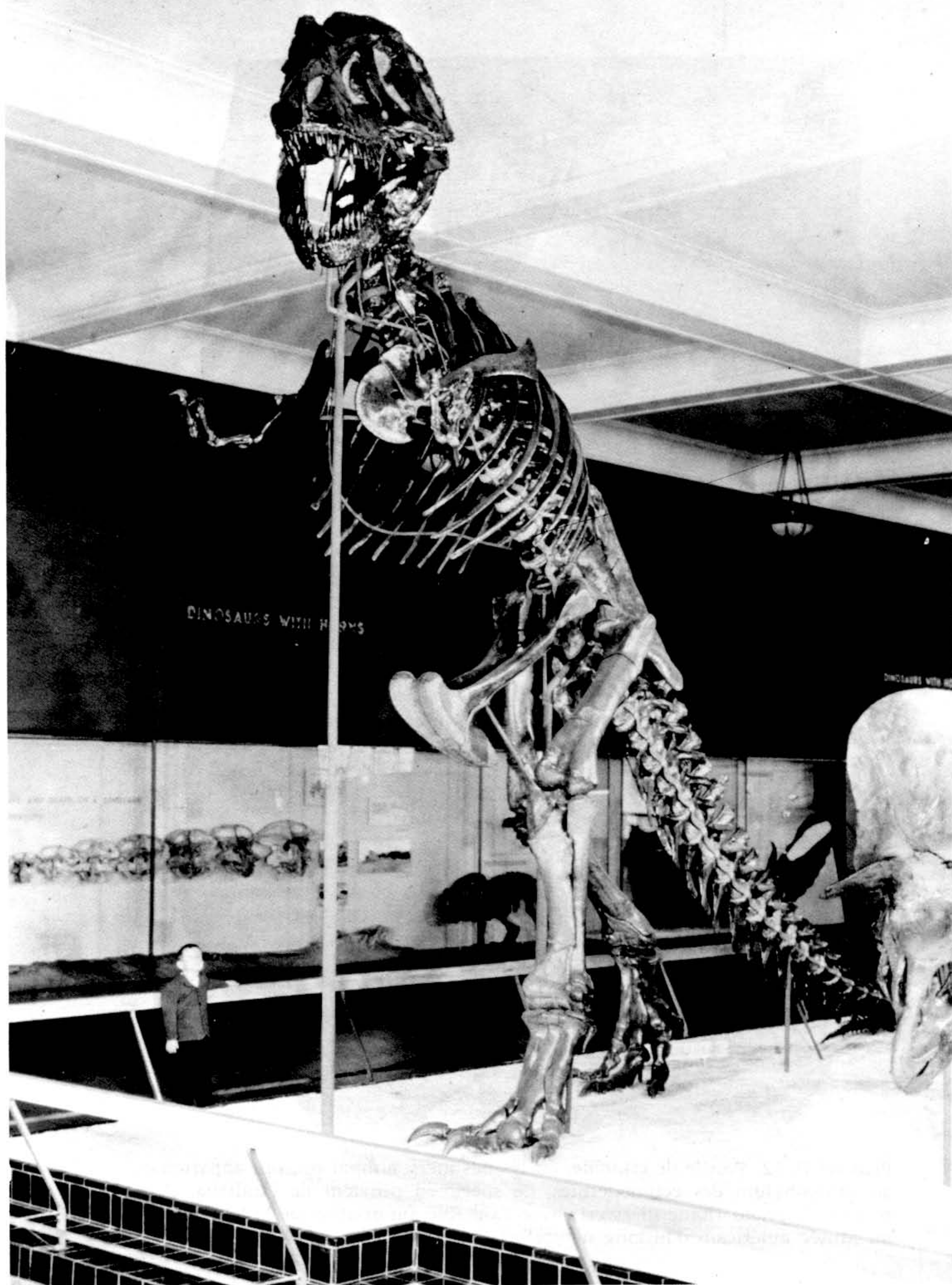


Planche IV.13. *Tyrannosaurus rex*, reconstruit à partir d'os fossilisés. On peut admirer ce spécimen dans la salle du Crétacé du Musée américain d'histoire naturelle de New York. Ce grand carnivore semait la terreur chez les dinosaures, des végétariens. (Avec l'aimable autorisation de la bibliothèque du Musée américain d'histoire naturelle.)

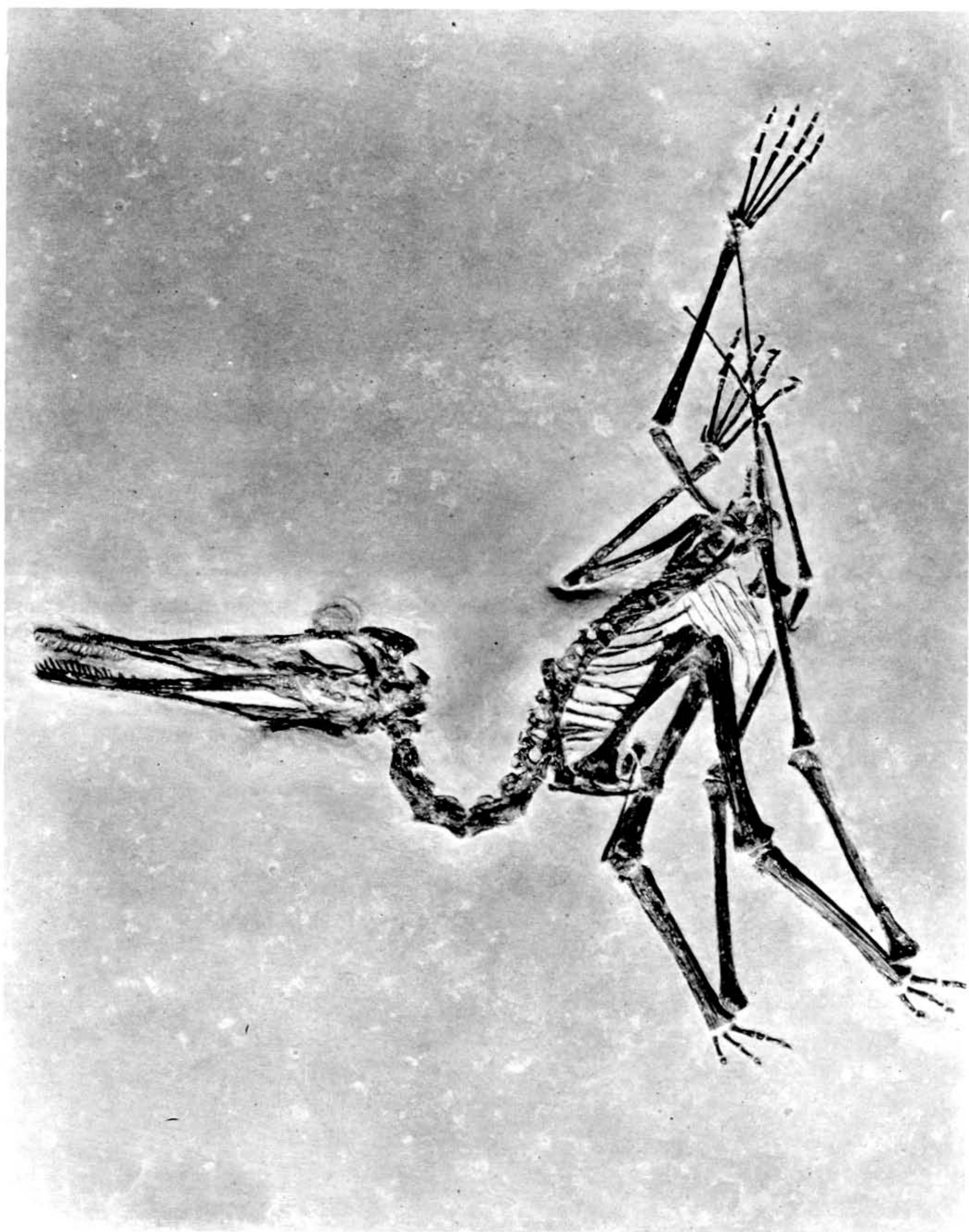
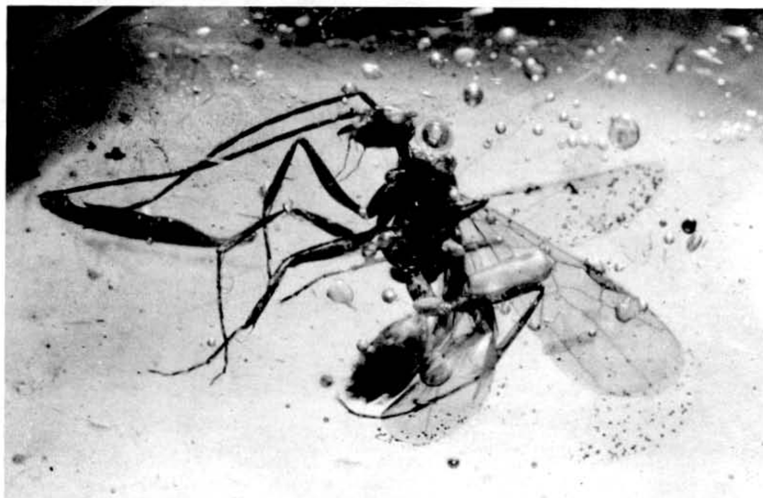


Planche IV.14. Squelette de ptérodactyle, un reptile ailé. (Nég. n° 315134, photo Charles H. Coles et Thane Bierwert ; avec l'aimable autorisation de la bibliothèque du Musée américain d'histoire naturelle.)

Planche IV.15. Moule d'un cœlacanthe. Ce poisson fossile vit encore dans les eaux profondes près de Madagascar. (Avec l'aimable autorisation de la bibliothèque du Musée américain d'histoire naturelle.)

Planche IV.16. Fourmi fossile délicatement préservée dans de l'ambre. (Avec l'aimable autorisation de la bibliothèque du Musée américain d'histoire naturelle.)



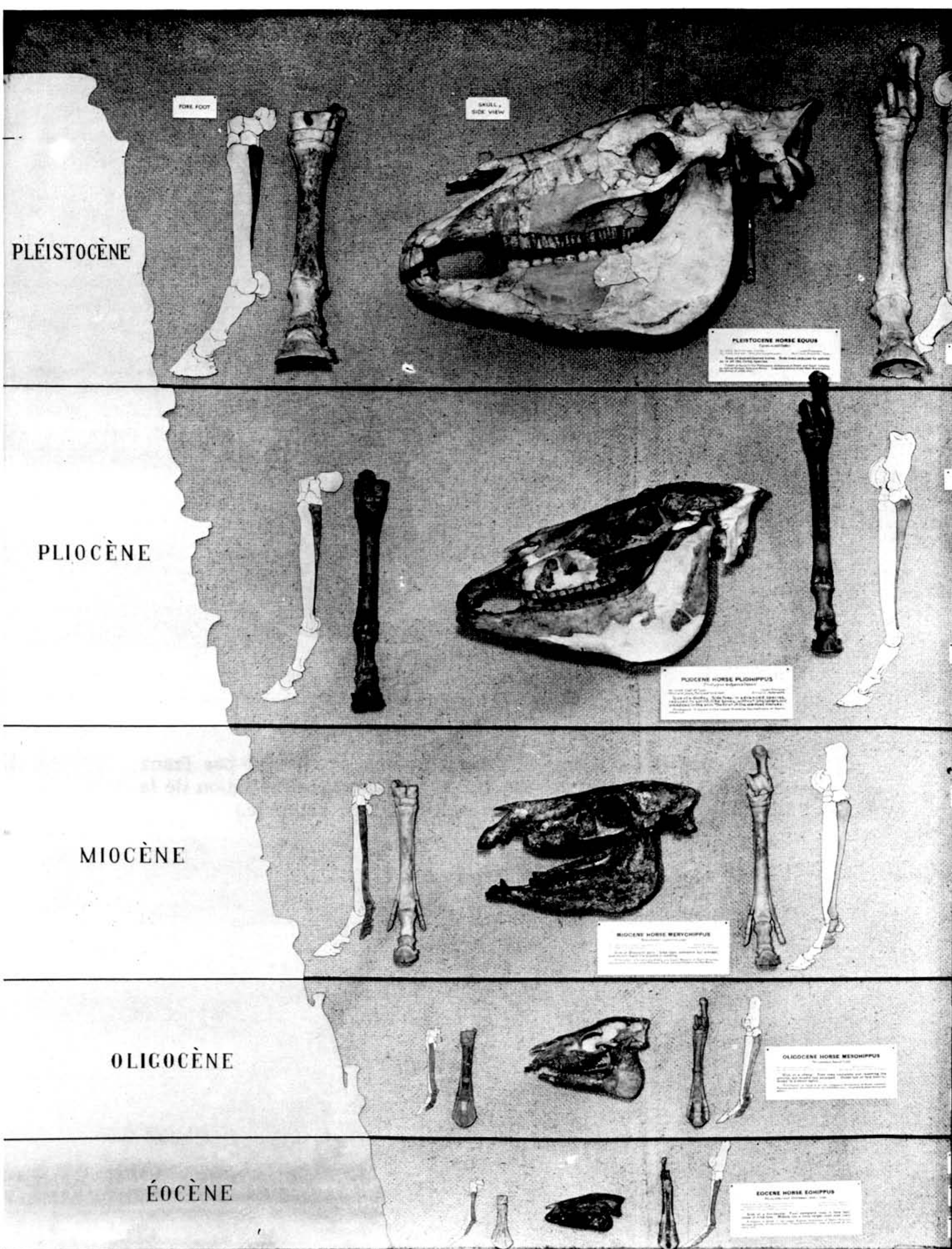


Planche IV.17. L'évolution du cheval, illustrée par le crâne et les os de la patte. (Nég. n° 322448, photo Baltini ; avec l'aimable autorisation de la bibliothèque du Musée américain d'histoire naturelle.)



Planche IV.18. Crâne de pithécanthrope, reconstruit par Franz Weidenreich. (Nég. n° 120979 ; avec l'aimable autorisation de la bibliothèque du Musée américain d'histoire naturelle.)

Planche IV.19. Femme sinanthrope, reconstruite par Franz Weidenreich et Lucile Swan. (Nég. nº 322021 ; avec l'aimable autorisation de la bibliothèque du Musée américain d'histoire naturelle.)





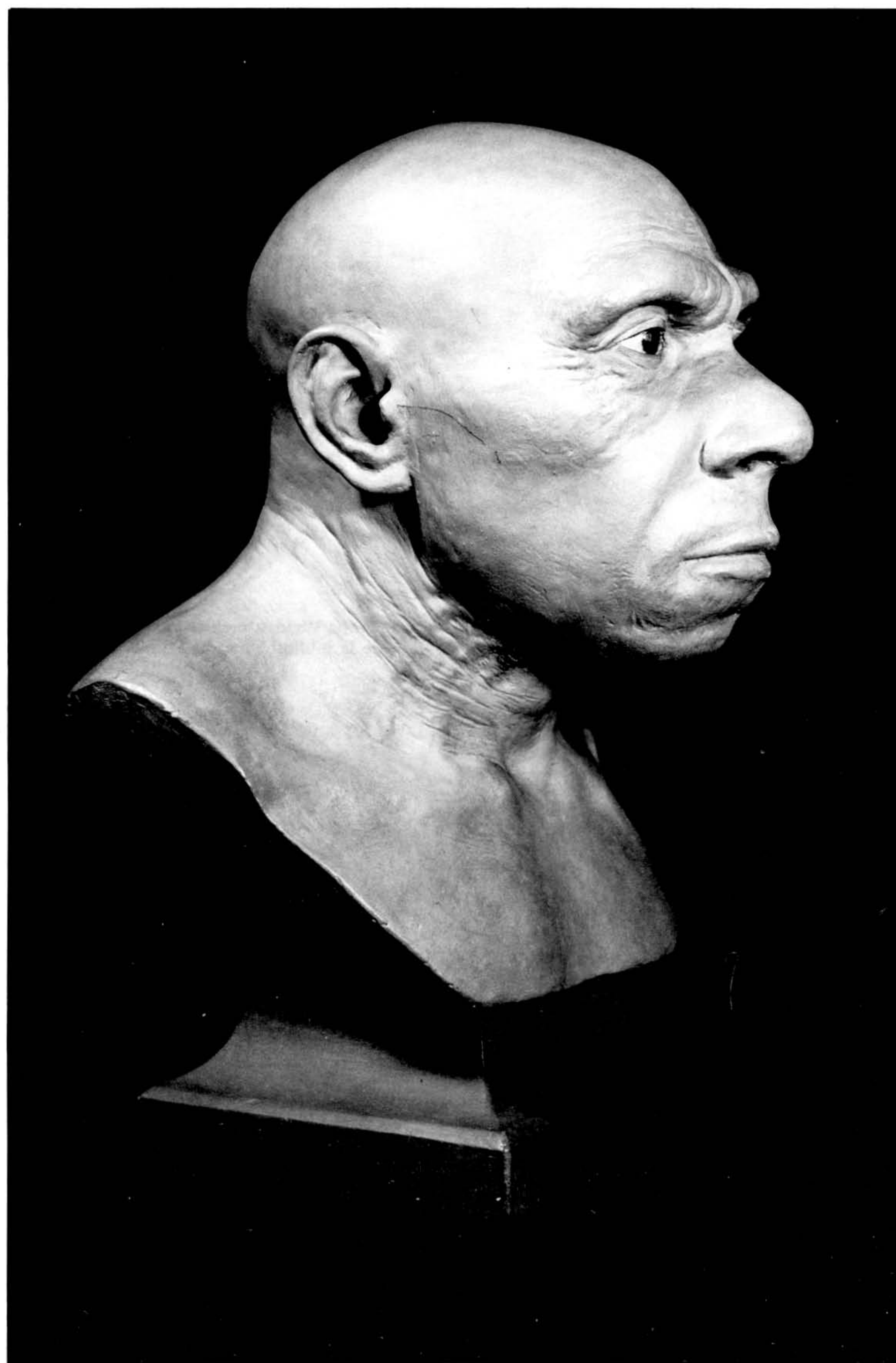


Planche IV.20. L'homme de Néanderthal, d'après un travail de restauration par J.H. McGregor. (Nég. n° 319951, photo Alex J. Rota ; avec l'aimable autorisation de la bibliothèque du Musée américain d'histoire naturelle.)

Planche IV.21. L'Univac des débuts, le premier gros ordinateur. (Avec l'aimable autorisation de Remington Rand Univac.)



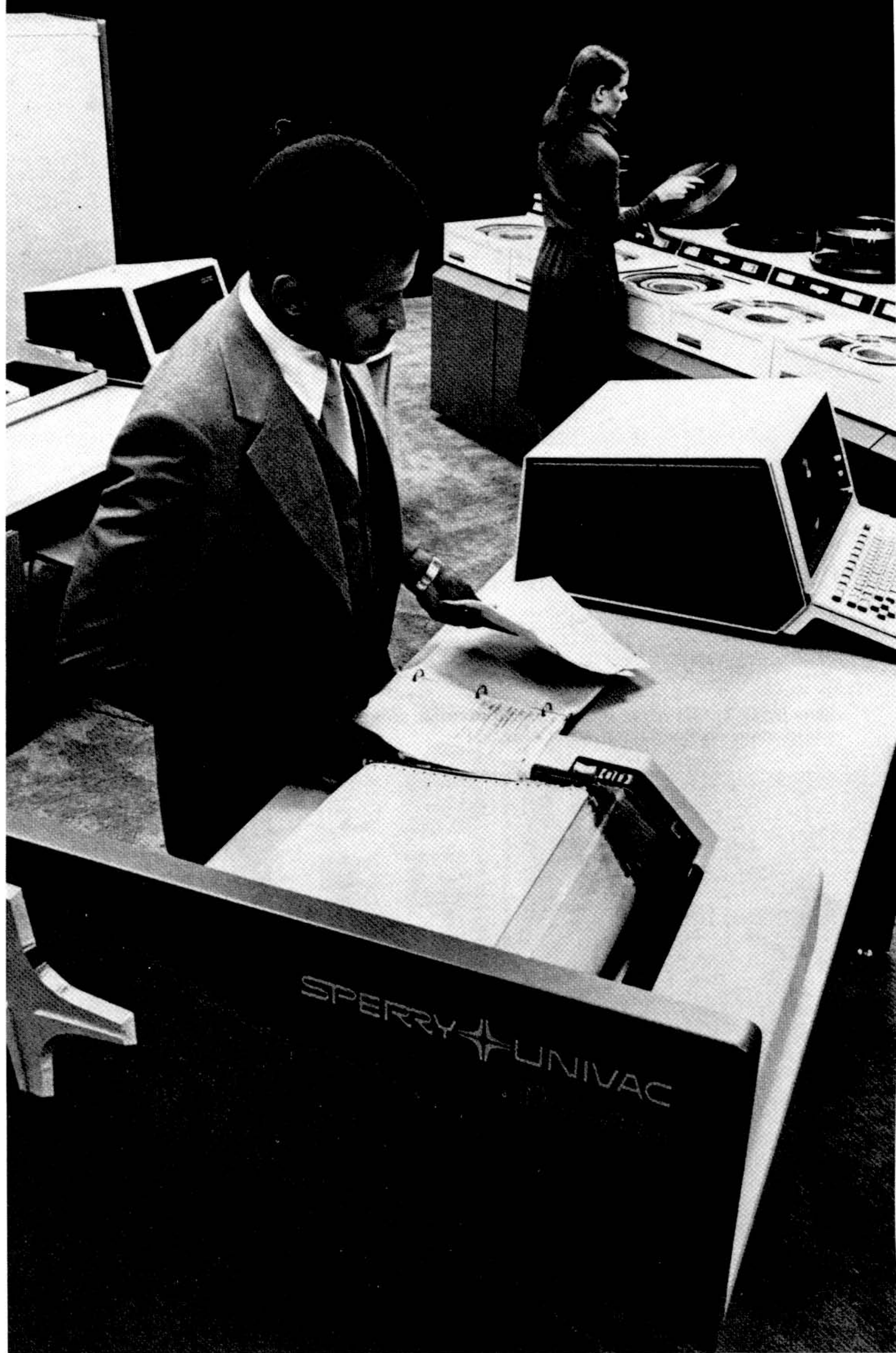
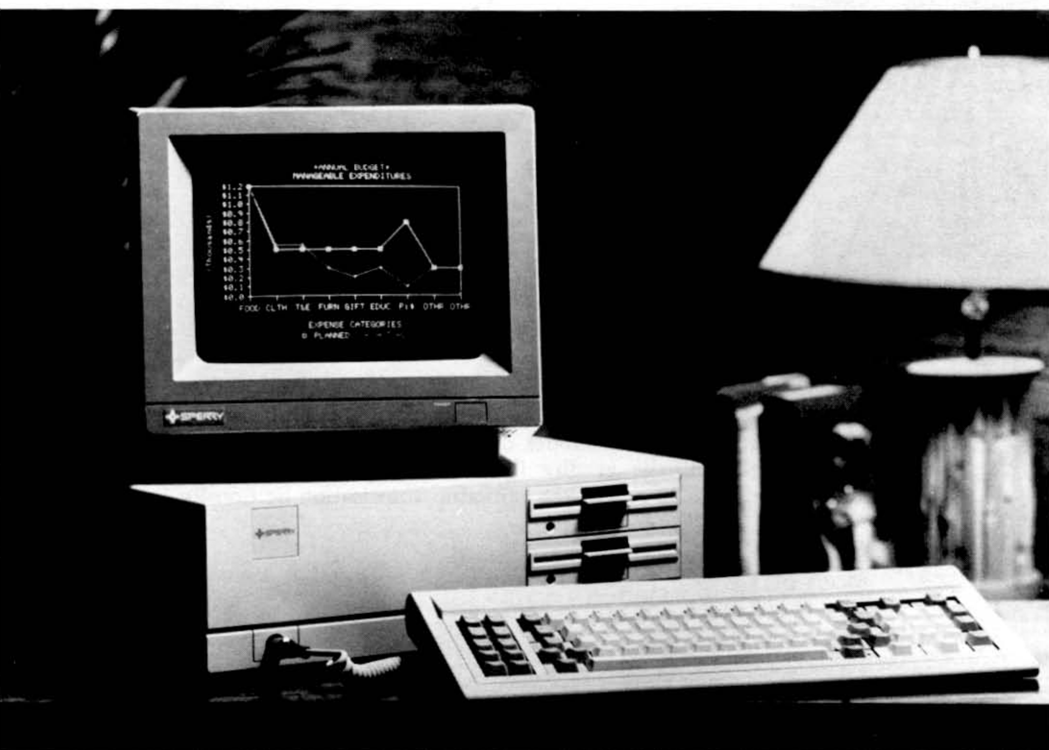

SPERRY UNIVAC

Planche IV.22. L'ordinateur commandé récemment par l'US Air Force. Grâce à leurs gigantesques bases de données et leurs vastes réseaux de communications, ces ordinateurs répondront aux besoins opérationnels et logistiques (pièces détachées d'avions, maintenance, etc.) de l'US Air Force dans le monde entier. En plus, ces nouveaux systèmes à base de microprocesseurs assureront un vaste éventail de fonctions administratives (gestion du personnel, comptabilité, génie civil, etc.). (Avec l'aimable autorisation de Sperry Corporation.)

Planche IV.23. Un ordinateur personnel. Celui-ci comporte une unité centrale, un moniteur et un clavier. (Avec l'aimable autorisation de Sperry Corporation.)



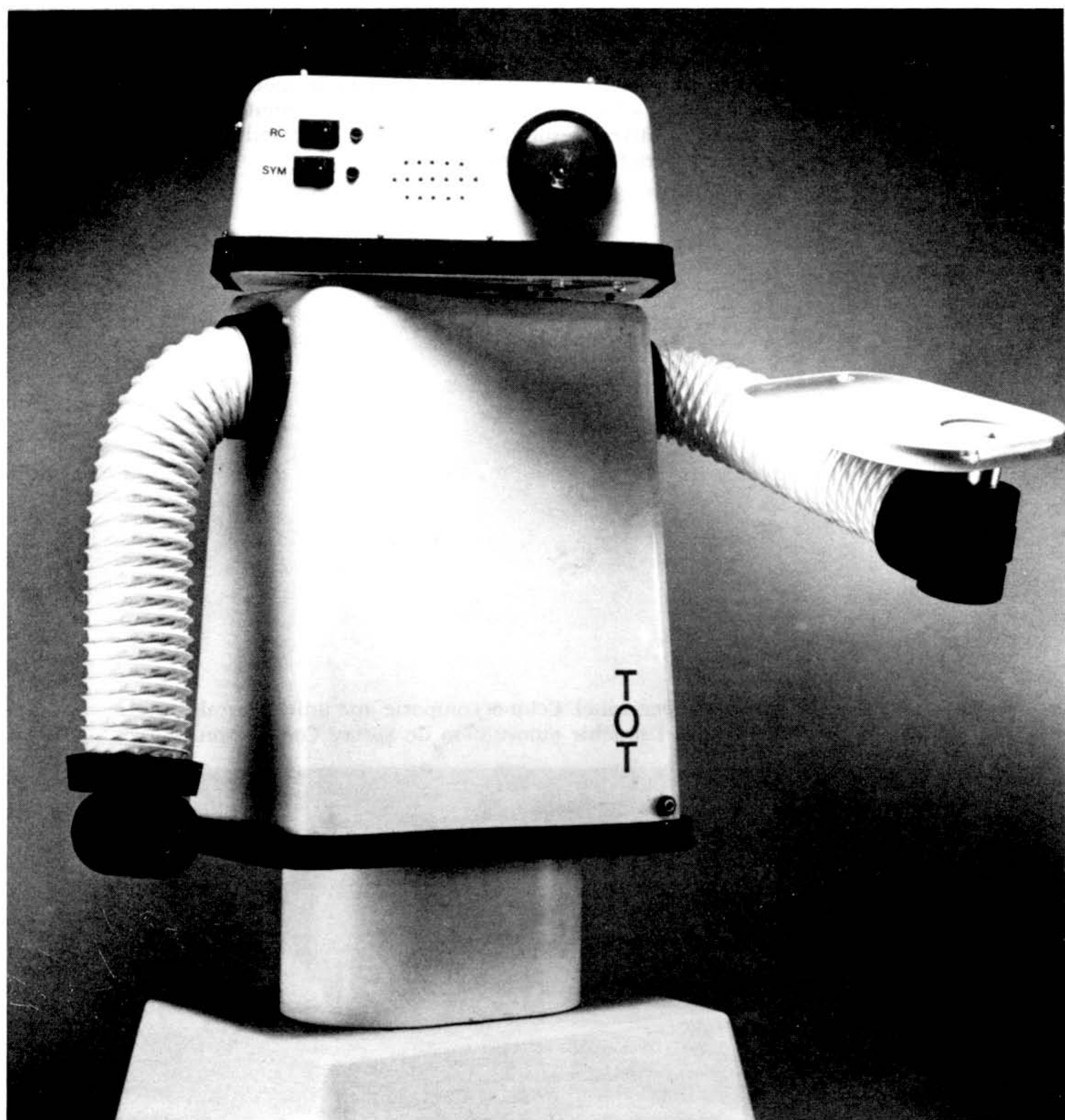


Planche IV.24. *Tot,* 1982-83. Ce robot personnel, mobile, programmable et multilingue est équipé de deux modes de contrôle et d'un système sensoriel ; il peut assurer le gardiennage et dire l'heure. Conçu par Jérome Hamlin ; ses dimensions : 90 × 60 × 30 cm. (Avec l'aimable autorisation de ComRo Inc., New York.)

Chapitre 16

Les espèces

La variété dans les formes de vie

Les connaissances que nous avons acquises sur notre corps seraient incomplètes si nous ignorions la question de notre interrelation avec les autres formes de vie qui existent sur Terre.

Dans beaucoup de cultures primitives, cette interrelation était des plus étroites. Certaines tribus considéraient tel ou tel animal comme leur ancêtre ou leur frère de sang, et allaient même jusqu'à interdire sa chasse ou sa consommation, sauf à l'occasion de certaines cérémonies rituelles. Cette vénération pour les animaux porte le nom de *totémisme* (d'un mot indien d'Amérique du Nord), et l'on en trouve des traces dans des cultures que l'on ne pourrait pourtant pas taxer de primitives. Les dieux à tête d'animaux chez les Égyptiens représentent un vestige du totémisme, comme peut-être la vénération actuelle que vouent les Hindous aux vaches et aux singes.

Au contraire, la culture occidentale, telle qu'elle s'exprime à travers les auteurs grecs et hébreux, établit très tôt une distinction marquée entre les êtres humains et les animaux « inférieurs ». C'est ainsi que la Bible souligne qu'Adam est né à l'image de Dieu (Genèse, I, 26) par un acte spécial de création. Pourtant, la Bible atteste de l'intérêt prononcé que l'homme porte très tôt aux animaux. Le livre de la Genèse raconte qu'Adam, pendant les premiers jours idylliques qu'il passe dans le jardin d'Eden, se voit confier par Dieu la tâche de donner un nom à « tous les animaux domestiques, toutes les bêtes sur la terre, à tous les oiseaux dans les cieux, et à toutes les bêtes des champs ».

A première vue, la tâche ne paraît pas insurmontable, l'affaire d'une heure ou deux tout au plus. Les auteurs des textes sacrés ont placé deux

représentants de chaque espèce dans l'arche de Noé, qui mesurait 150 mètres de long, 25 de large et 15 de haut. Les philosophes naturalistes grecs se représentaient le monde vivant en des termes aussi limités : Aristote a établi une liste qui ne comportait que cinq cents espèces animales, et son élève Théophraste, le botaniste le plus réputé de la Grèce antique, n'a recensé qu'environ cinq cents plantes différentes.

Mais une telle liste n'a de sens que si l'on postule que l'éléphant n'a été, n'est et ne sera toujours qu'un éléphant, un chameau toujours un chameau, et une puce rien, jamais rien d'autre qu'une puce. La situation s'est singulièrement compliquée lorsque les naturalistes se sont aperçus qu'on pouvait différencier les animaux par leur aptitude (ou inaptitude) à se reproduire entre eux. L'éléphant des Indes ne peut se reproduire avec l'éléphant d'Afrique : il faut donc les considérer comme des *espèces* différentes d'éléphant ; de même, le chameau d'Arabie (à une bosse) et le chameau d'Asie (à deux bosses) sont deux espèces différentes. Et pour ce qui est de la puce, on ne compte pas moins de cinq cents espèces différentes de ce petit insecte vorace, toutes très semblables à la puce commune.

Au cours des siècles, au fur et à mesure que les naturalistes ont recensé de nouvelles formes de vie sur terre, dans les airs et sous la mer, et grâce à la découverte de nouvelles contrées par les explorateurs, le nombre d'espèces animales et végétales identifiées a augmenté de façon vertigineuse. En 1800, ce chiffre était de 70 000. Aujourd'hui, on connaît plus d'un million et demi d'espèces différentes – deux tiers animales et un tiers végétales –, et aucun biologiste ne pense qu'il s'agisse là d'une liste exhaustive.

Il reste même probablement à découvrir des animaux de taille respectable dans certaines régions reculées de notre globe. L'okapi, un cousin de la girafe mais de la taille d'un zèbre, n'a été connu des biologistes qu'en 1900 lorsqu'on a enfin pu le capturer dans les forêts du Congo. Aussi récemment qu'en 1983, une nouvelle espèce d'albatros a été découverte sur une île de l'océan Indien, ainsi que deux espèces de singe dans les jungles d'Amazonie.

Il ne fait aucun doute qu'un grand nombre d'espèces inconnues demeurent cachées au plus profond des océans, car leur exploration systématique y demeure encore à ce jour particulièrement difficile. L'existence du calmar géant, le plus grand de tous les invertébrés, n'a été révélée que dans les années 1860. Le *coelacantha ornata* (voir chapitre 4) n'a été découvert qu'en 1938.

Et pour ce qui est des petits animaux – insectes, lombrics, etc. – on en découvre de nouvelles variétés tous les jours. Une évaluation modeste donne à penser qu'il existe actuellement dix millions d'espèces sur Terre. S'il est exact qu'environ les neuf dixièmes des espèces ayant jamais existé ont aujourd'hui disparu, ce sont cent millions d'espèces qui ont vécu sur le globe à un moment ou à un autre.

CLASSIFICATION

La plus grande confusion régnerait dans nos connaissances sur le monde vivant si nous ne parvenions pas à mettre un peu d'ordre dans cette immense variété de créatures, et ne suivions un quelconque schéma logique. On peut par exemple commencer par ranger ensemble le chat, le tigre, le lion, la panthère, le léopard, le jaguar, et d'autres animaux ressemblant au chat dans la *famille des chats ;* de la même manière, on peut classer le chien, le loup, le renard, le chacal et le coyote dans une

famille des chiens, etc. En se basant sur des critères généraux évidents, on peut classer certains animaux comme carnivores et d'autres comme herbivores. Les Anciens avaient également établi des classifications générales basées sur l'habitat, et considéraient ainsi tous les animaux qui vivent dans la mer comme des poissons et tous ceux qui volent comme des oiseaux. Mais si l'on suit ce raisonnement, la baleine devient un poisson et la chauve-souris un oiseau. En fait, fondamentalement, la baleine et la chauve-souris se ressemblent plus que l'une ne ressemble à un poisson et l'autre à un oiseau. Les deux mettent au monde des petits qui sont déjà formés et vivants à la naissance. De plus, la baleine possède des poumons pour respirer, plutôt que des ouïes comme c'est le cas chez les poissons, et la chauve-souris a des poils et non des plumes comme les oiseaux. Toutes deux sont classées parmi les mammifères, qui mettent au monde des petits vivants (au lieu de pondre des œufs), et les allaitent.

Une des premières tentatives de classification systématique revient à un Anglais, John Ray (ou Wray), qui au XVIIe siècle classifia toutes les espèces végétales connues (soit environ 18 600), et plus tard les espèces animales, selon des schémas qui lui semblaient logiques. Par exemple, il divisa les plantes à fleurs en deux groupes principaux, en se basant sur la présence d'une ou deux feuilles embryonnaires dans la graine. Il baptisa la minuscule feuille *cotylédon,* d'un mot grec signifiant « creux d'une coupe » *(kotulêdon),* car elle se trouvait placée au fond de la graine, dans une cavité ressemblant à une coupe. C'est ainsi que Ray donna respectivement aux deux types de plantes les noms de *monocotylédone* et *dicotylédone.* Cette classification (similaire, notons-le, à celle établie par Théophraste deux mille ans auparavant) s'est révélée si utile qu'elle est encore utilisée aujourd'hui. La différence entre une feuille embryonnaire et deux est en elle-même sans grande importance, mais il est certain qu'un grand nombre de différences distinguent les plantes monocotylédones des dicotylédones. Il s'avère que le nombre de feuilles embryonnaires est une étiquette commode, et qui se révèle symptomatique d'un grand nombre d'autres différences. Dans un même ordre d'idées, la distinction faite entre plumes et poils est mineure en elle-même, mais plumes et poils sont effectivement des marqueurs aisément identifiables des innombrables différences entre les oiseaux et les mammifères.

Bien que Ray et d'autres aient apporté quelques idées très utiles, le véritable fondateur des lois de la classification, la *taxonomie* (d'un mot grec signifiant « organisation »), fut le botaniste suédois Carl von Linné, dont le travail fut d'une telle qualité que les grandes lignes de son système sont encore valides de nos jours. Linné présenta sa méthode dans un livre paru en 1737 sous le titre *Systema Naturae.* Il regroupa les espèces se ressemblant dans un même *genre* (du latin *genus* pour « origine, naissance »), les genres reliés entre eux dans un *ordre,* et les ordres similaires dans une *classe.* Il donna à chaque espèce un double nom, construit à partir du nom du genre et de celui de l'espèce proprement dite. Ce système rappelle celui de l'annuaire téléphonique, dans lequel on trouve les Dupont (Jean Dupont, Pierre, etc.). C'est ainsi que les membres du genre des chats sont *Felix domesticus* (le chat domestique), *Felix leo* (le lion), *Felix tigris* (le tigre), *Felix pardus* (le léopard), etc. Le genre auquel appartient le chien comprend *Canis familiaris* (le chien), *Canis lupus* (le loup gris d'Europe), *Canis occidentalis* (le loup d'Amérique), etc. Les deux espèces de chameau sont *Camelus bactrianus* (le chameau d'Asie) et *Camelus dromaderius* (le chameau d'Arabie).

Vers 1800, le naturaliste français Georges Léopold Cuvier ajouta à la notion de *classe* une catégorie plus générale qu'il appela *phylum* (du grec *phulon* : « race », « tribu »). Un phylum, ou encore embranchement, comprend tous les animaux partageant un plan corporel général similaire (un concept mis en valeur et clarifié par nul autre que le grand poète allemand Johann Wolfgang von Goethe). Par exemple, mammifères, oiseaux, reptiles, amphibiens et poissons sont classés dans le même phylum parce qu'ils ont tous une épine dorsale, un maximum de quatre membres, et du sang rouge contenant de l'hémoglobine. Insectes, araignées, homards et scolopendres (mille-pattes) sont dans un autre embranchement ; les palourdes, huîtres et moules dans un autre encore, etc. Dans les années 1820, le botaniste suisse Augustin Pyramus de Candolle améliora de la même manière la classification des plantes de Linné. Au lieu de regrouper les espèces selon leur apparence externe, il accorda plus de signification à la structure et au fonctionnement internes.

Je décrirai dans les paragraphes qui suivent l'arbre de vie tel qu'il est aujourd'hui organisé, en allant des divisions les plus générales vers les plus spécifiques.

Nous débuterons par les *règnes*, que l'on a longtemps fixés au nombre de deux, l'animal et le végétal. Néanmoins, le développement de nos connaissances sur les micro-organismes a quelque peu compliqué les choses ; c'est ainsi que le biologiste américain Robert Harding Whittaker a suggéré qu'on subdivise les micro-organismes vivants en au moins cinq règnes.

Le système de Whittaker repose sur l'idée qu'on doit réserver les règnes animal et végétal aux organismes pluricellulaires. Les plantes sont caractérisées par le fait qu'elles possèdent de la chlorophylle (c'est pour cela qu'on les appelle communément les *plantes vertes*) et qu'elles sont le siège d'un processus complexe, la photosynthèse. Les animaux, eux, mangent d'autres organismes et possèdent un système digestif.

Un troisième règne, les *moisissures*, se compose d'organismes pluricellulaires qui ressemblent par certains points aux plantes, mais qui sont dépourvus de chlorophylle. Ils vivent aux dépens d'autres organismes, mais ne les ingèrent pas comme les animaux ; ils excrètent des enzymes de digestion, digèrent leur nourriture hors de l'organisme, puis l'absorbent.

Les deux derniers règnes comprennent les organismes unicellulaires. Les *protistes*, un terme inventé en 1866 par le biologiste allemand Ernst Heinrich Haeckel, incluent les eucaryotes : ces derniers sont subdivisés en cellules ressemblant à celles des animaux (les *protozoaires*, comme l'amibe et la paramécie), et à celles des plantes (les *algues*).

Enfin, le règne des *monères* comprend les organismes unicellulaires qui sont procaryotes – les bactéries et l'algue bleu-vert. Délaissés de ce système sont les virus et les viroïdes, des organismes sous-cellulaires, qui peut-être forment un sixième règne.

Le règne végétal, selon un système de classification, se divise en deux phylums majeurs – les bryophytes (les diverses mousses) et les trachéophytes (les plantes possédant un réseau de tubes permettant la circulation de la sève), qui comprennent toutes les espèces que l'on range communément sous le vocable de plantes.

Ce dernier grand phylum se subdivise en trois classes : les filinicées, les gymnospermes, et les angiospermes. Dans la première classe, on trouve les fougères, qui se reproduisent grâce à des spores. Les gymnospermes, qui forment des graines à la surface d'organes spécialisés, incluent les divers

conifères à feuilles non caduques. Les angiospermes, dont les graines sont contenues dans des ovules, comprennent la grande majorité de nos plantes familières.

Pour ce qui est du règne animal, je ne mentionnerai que les phylums les plus importants.

Les *porifères* sont des animaux constitués par des colonies de cellules à l'intérieur d'un squelette formé de pores : ce sont les éponges. Les cellules individuelles manifestent des signes de spécialisation, mais conservent une certaine indépendance, puisque si on les sépare en les filtrant au travers d'un morceau de soie, elles peuvent se rassembler pour former une nouvelle éponge.

(D'une façon générale, au fur et à mesure que les phylums animaux deviennent plus spécialisés, les cellules et les tissus individuels deviennent moins « indépendants ». Les animaux peu évolués peuvent retrouver leur intégrité d'origine même s'ils sont gravement mutilés, un processus qu'on appelle la *régénération.* Des animaux plus complexes peuvent régénérer de nouveaux membres. Chez l'homme, en revanche, cette capacité n'est que très limitée. Un ongle perdu parvient à repousser, pas un doigt.)

Les coelenterata (« intestins vides ») forment le premier phylum dont les membres peuvent être considérés comme de véritables animaux pluricellulaires. Ces animaux ont en gros la forme d'une coupe et sont constitués de deux couches de cellules – l'*ectoderme* (« peau externe ») et l'*endoderme* (« peau interne »). Les exemples les plus communs de ce phylum sont la méduse et l'anémone de mer.

Tous les autres phylums animaux possèdent une troisième couche de cellules – le *mésoderme* (« peau intermédiaire »). A partir de ces trois couches, mises en évidence en 1845 par les physiologistes allemands Johannes Müller et Robert Remak, se forment les nombreux organes existant chez les animaux les plus développés, dont l'homme.

Le mésoderme apparaît pendant le développement de l'embryon, et la manière dont il se forme donne lieu à une division en deux *superphylums* chez les animaux concernés. Ceux chez qui le mésoderme se forme à la jonction entre l'ectoderme et l'endoderme constituent le superphylum des annélides ; ceux chez qui le mésoderme provient du seul endoderme forment le superphylum des échinodermes.

Considérons tout d'abord les annélides. L'embranchement le plus simple est celui des plathelminthes (du grec signifiant « ver plat »). Sont inclus dans ce phylum non seulement le ver solitaire, mais également d'autres formes de vies autonomes. Les vers possèdent des fibres contractiles semblables à des muscles primitifs ; ils ont aussi une tête, une queue, des organes de reproduction spécialisés, et les rudiments d'organes d'excrétion. De plus, les vers se caractérisent par une symétrie bilatérale : ils ont un côté gauche qui est l'image inversée du côté droit. Ils se déplacent « la tête la première » ; leurs organes sensoriels ainsi que leur système nerveux rudimentaire sont concentrés dans cette zone, si bien que l'on peut dire que les vers plats possèdent les rudiments d'un cerveau.

Puis vient le phylum des nématodes, dont le membre le plus connu est l'ankylostome. Ces animaux ont une circulation sanguine primitive : un liquide à l'intérieur du mésoderme baigne toutes les cellules et transporte vers elles la nourriture et l'oxygène. Cela donne aux nématodes une certaine épaisseur, à l'inverse du ver solitaire plat, car ce liquide apporte des aliments aux cellules internes. Les nématodes possèdent

également un boyau qui comporte deux ouvertures, une pour l'ingestion des aliments, l'autre (l'anus) pour l'éjection des déchets.

Les deux phylums suivants de ce superphylum sont dotés d'un *squelette* externe – c'est-à-dire d'une coquille (qui existe également chez certains phylums plus primitifs). Il s'agit des lophophoriens, qui possèdent une coquille formée de carbonate de calcium sur le dessus et le dessous, et qu'on appelle communément vers à lophophore en « fer à cheval », et des mollusques (du latin « mou »), dont le corps mou est enveloppé d'une carapace dont les origines se trouvent sur les côtés droit et gauche au lieu de dessus et dessous. Les mollusques les plus connus sont la palourde, l'huître et l'escargot.

Un embranchement particulièrement important dans le superphylum des annélides est *annelida*. Il s'agit de vers, mais avec une différence : ils sont composés de segments, dont chacun peut être considéré comme un organisme en soi. Chaque segment possède ses propres nerfs qui se ramifient à partir de l'axe nerveux principal, ses propres vaisseaux sanguins, ses propres canaux pour le transport des déchets, ses propres muscles, etc. Chez le plus connu des annélides, le ver de terre, les segments sont marqués par de petites constrictions de la peau qui ressemblent à des anneaux autour de l'animal ; de fait, le mot *annelida* vient du latin « petit anneau ».

Cette segmentation semble procurer à l'animal une plus grande efficacité, car toutes les espèces du règne animal qui se sont le mieux adaptées, dont l'espèce humaine, sont segmentées. Le plus complexe et le mieux adapté des animaux non segmentés est le calmar. Chez l'homme, les vertèbres et les côtes sont des exemples parmi tant d'autres de ce phénomène ; chaque vertèbre de la colonne vertébrale et chaque côte représentent des segments uniques du corps, dotés de nerfs, muscles et vaisseaux sanguins propres.

Les annélides sont dépourvus de squelette, et sont donc mous et pratiquement sans défense. L'embranchement des arthropodes (« pieds joints »), cependant, associe segmentation et squelette, ce dernier étant segmenté comme le reste du corps. Le squelette n'est pas seulement plus « manœuvrable » grâce à ses articulations, il est également léger et solide, car constitué d'un polysaccharide (la *chitine*) et non pas de calcaire ou de carbonate de calcium, des substances lourdes et rigides. Dans l'ensemble, les arthropodes – dont font partie le homard, les araignées, les scolopendres et les insectes – constituent le phylum qui, à ce jour, a le mieux réussi. Ou tout au moins comporte-t-il plus d'espèces que tous les autres phylums réunis.

Voilà donc pour ce qui est du phylum principal dans le superphylum des annélides. L'autre superphylum, les échinodermes, ne comporte que deux grands embranchements. L'un s'appelle *echinodermata* (« peau couverte d'épines »), et comprend des animaux tels que l'étoile de mer et l'oursin. Les échinodermes se différencient des autres phylums possédant un mésoderme par leur symétrie rayonnante et par l'absence de tête ou de queue clairement marquées (bien que, dans le premier stade de leur développement, les échinodermes manifestent une symétrie bilatérale, qu'ils perdent au cours de leur croissance).

Le second grand phylum des échinodermes revêt une importance toute particulière, puisque c'est celui auquel appartient notre espèce.

LES VERTÉBRÉS

Le squelette interne (figure 16.1) représente la marque distinctive des membres de ce phylum (dont font partie l'homme, l'autruche, le serpent,

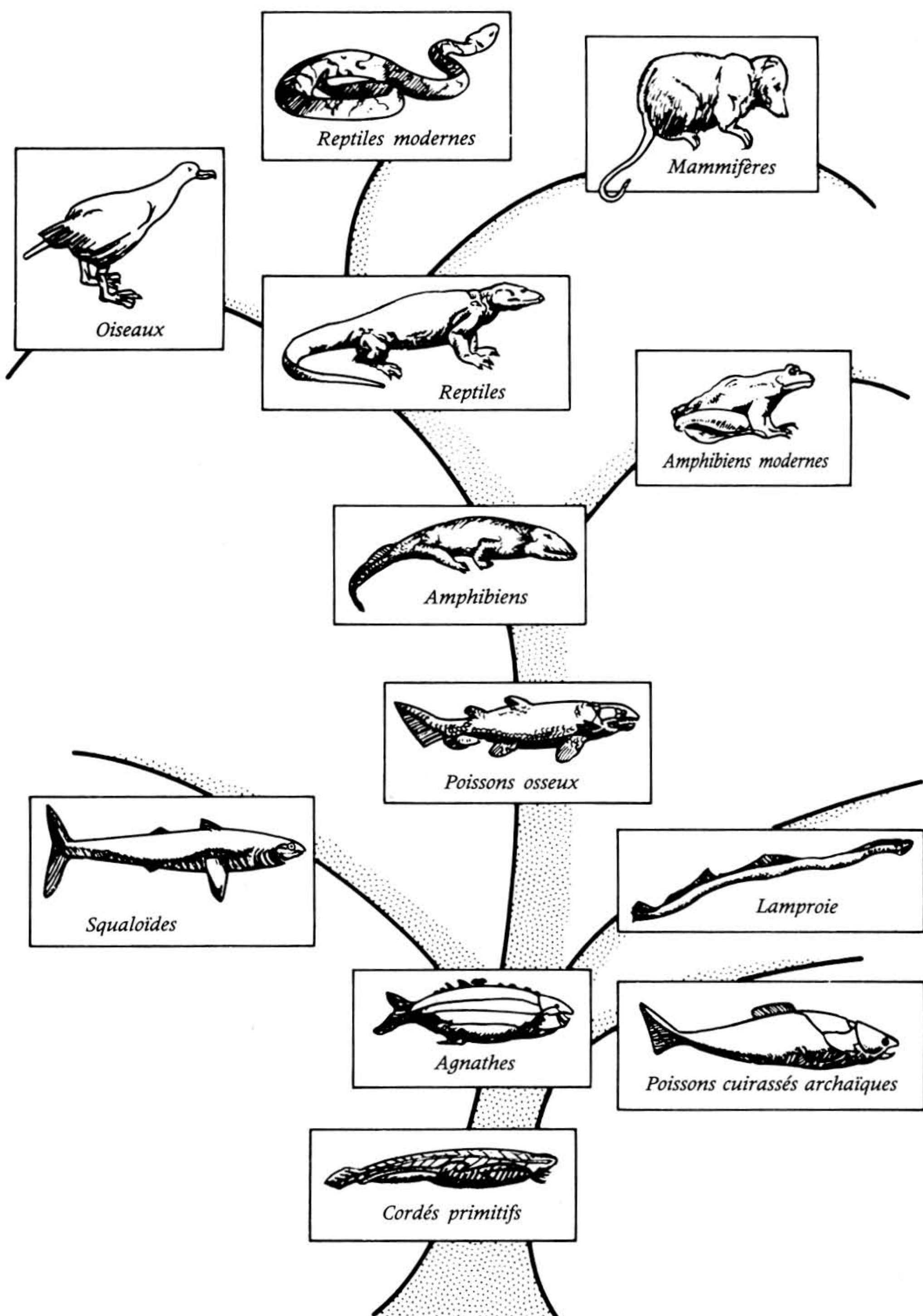


Figure 16.1. Un arbre philogénétique, montrant l'évolution des vertébrés.

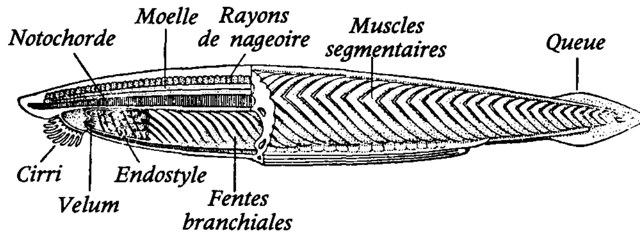


Figure 16.2. L'amphioxus, un cordé primitif avec sa corde dorsale.

la grenouille, le maquereau, ainsi qu'une foule d'autres animaux). Aucun autre animal en dehors de cet embranchement n'en possède un. C'est l'épine dorsale qui caractérise ce type de squelette. En fait, l'épine dorsale est une caractéristique si importante que le langage ordinaire a pris l'habitude de diviser les animaux entre vertébrés et invertébrés. En réalité, il existe un groupe intermédiaire qui se distingue par une tige cartilagineuse appelée *corde dorsale* à la place de la colonne vertébrale (figure 16.2). La corde dorsale, mise en évidence par Von Baer, à qui l'on doit également la découverte de l'ovule chez les mammifères, ressemble à une colonne vertébrale rudimentaire ; de fait, elle apparaît même chez les mammifères au cours du développement de l'embryon. C'est pourquoi les animaux qui possèdent une corde dorsale (c'est-à-dire certains animaux ressemblant au ver, à la limace et à divers mollusques) sont classés parmi les vertébrés. Cet embranchement a été baptisé *cordés* en 1880 par le zoologiste anglais Francis Maitland Balfour ; il comprend quatre subphylums, dont trois seulement possèdent une corde dorsale. Le quatrième, caractérisé par une véritable colonne vertébrale et un squelette interne, s'appelle *vertebrata*.

Les vertébrés d'aujourd'hui se divisent en deux superclasses : les poissons et les tétrapodes (animaux à « quatre pattes »).

Le groupe des poissons comprend trois classes : les poissons agnathes (« sans mâchoires »), qui possèdent un vrai squelette mais sont dépourvus de membres et de mâchoires – le représentant le plus connu, la lamproie, possède une bouche ronde en forme de suçoir, remplie de petites dents râpeuses ressemblant à des limes ; les chondrichthyens (poissons à cartilage), dotés d'un squelette cartilagineux et non pas osseux, dont les requins sont les exemples les plus typiques ; enfin les ostéichthyens, ou poissons à arêtes.

Les tétrapodes respirent tous au moyen de poumons ; ils sont divisés en quatre classes. Les plus primitifs sont les amphibiens (« double vie ») – par exemple les grenouilles et les crapauds. Ils mènent effectivement une « double vie » : jeunes (les têtards), ils sont dépourvus de membres et respirent à l'aide de branchies ; une fois adultes, ils ont quatre membres et des poumons. Les amphibiens, comme les poissons, pondent leurs œufs dans l'eau.

Les reptiles (du latin « rampant ») forment la seconde classe. Ils comprennent les serpents, lézards, alligators et tortues. Ils respirent au moyen de poumons dès la naissance, et font éclore leurs œufs (protégés par une coquille dure) sur terre. Les reptiles les plus évolués possèdent

un cœur à quatre chambres, alors que celui des amphibiens en a trois, et celui des poissons seulement deux.

Les deux derniers groupes de tétrapodes sont les oiseaux et les mammifères. Tous sont caractérisés par un sang chaud, c'est-à-dire qu'ils maintiennent une température interne constante quelle que soit la température extérieure (cela, bien sûr, dans des limites raisonnables). Comme leur température interne est souvent plus élevée que celle qui règne à l'extérieur de leur corps, ces animaux ont besoin d'un « isolant ». A cet effet, les oiseaux possèdent des plumes et les mammifères des poils, les deux contribuant à la formation au contact de la peau d'une couche d'air isolante. Les oiseaux pondent des œufs tout comme les reptiles. Les mammifères, bien sûr, mettent au monde des petits déjà « éclos », et les nourrissent avec du lait sécrété par les glandes mammaires (en latin *mammae*).

Au XIXe siècle, les zoologistes ont appris une nouvelle tellement extraordinaire qu'ils ont tout d'abord refusé de la croire. Les Australiens avaient découvert un animal possédant des poils et allaitant ses petits (bien que les glandes mammaires soient dépourvues de mamelons), et qui pourtant pondait des œufs ! Même lorsqu'ils purent examiner des spécimens de l'animal (morts, hélas, car il était difficile de les maintenir en vie hors de leur habitat naturel), les zoologistes crurent tout d'abord qu'il s'agissait là d'un canular grossier. L'animal vivait dans l'eau et sur terre, et ressemblait beaucoup à un canard : il possédait un bec et des pattes palmées. Finalement, ils durent bien admettre l'existence de l'*ornithorynque*, un mammifère. L'échidné, un autre mammifère pondant des œufs, a été découvert depuis en Australie et en Nouvelle-Guinée. Mais ces mammifères ne se rapprochent pas seulement des reptiles parce qu'ils pondent des œufs. La régulation thermique de leur circulation sanguine est également imparfaite : lorsqu'il fait froid, leur température interne peut baisser de 10 °C.

Les mammifères se divisent en trois sous-classes. Ceux qui pondent des œufs forment la première, les protothériens (du grec pour « premiers animaux »). L'embryon contenu dans l'œuf est en fait à un stade avancé de son développement au moment de la ponte, et il éclôt peu de temps après. La seconde sous-classe des mammifères, les métathériens (« animaux moyens »), comprend l'opossum et le kangourou. Bien qu'ils naissent vivants, les petits sont à un stade encore peu développé et meurent rapidement s'ils ne parviennent pas à trouver le hâvre protecteur de la poche maternelle dans laquelle se trouvent les glandes mammaires, et qu'ils ne quitteront pas avant d'avoir acquis la force nécessaire pour affronter le monde extérieur. On appelle ces animaux les marsupiaux (du latin *marsupium*, « poche »).

Enfin, au sommet de la hiérarchie des mammifères, nous arrivons à la sous-classe des euthériens (« vrais animaux »). C'est le placenta qui les caractérise, un tissu irrigué par des vaisseaux sanguins qui permet à la mère de nourrir l'embryon, et d'assurer ainsi son développement pendant une longue période de gestation à l'intérieur de son corps (neuf mois chez l'être humain, deux ans chez l'éléphant et la baleine). On appelle ces animaux les *animaux placentaires*.

Ces derniers se divisent en une bonne douzaine d'ordres, comme par exemple :

- les insectivores (« mangeurs d'insectes ») : la musaraigne, la taupe, etc. ;

- les chiroptères (« mains en forme d'ailes ») : la chauve-souris ;
- les carnivores (« mangeurs de viande ») : la famille du chat, la famille du chien, la belette, l'ours, le phoque, etc., mais pas l'homme ;
- les rongeurs : la souris, le rat, le lapin, l'écureuil, le cochon d'Inde, le porc-épic, etc. ;
- les édentés : le paresseux et le tatou, qui possèdent des dents, et le fourmilier, qui n'en possède pas ;
- les artiodactyles (« doigts pairs ») : des animaux à sabots qui ont un nombre pair de doigts à chaque pied, comme la vache, le mouton, la chèvre, le porc, le cerf, l'antilope, le chameau, la girafe, etc. ;
- les périssodactyles (« doigts impairs ») : le cheval, l'âne, le zèbre, le rhinocéros et le tapir ;
- les proboscidés (« nez long ») : l'éléphant, bien sûr ;
- les odontocètes (« baleines à dents ») : le cachalot et d'autres possédant des dents ;
- les mysticètes (« baleines à moustaches ») : la baleine blanche, la baleine bleue, et d'autres qui filtrent leur nourriture au travers de lames cornées qui ressemblent à une immense moustache à l'intérieur de leur gueule ;
- les primates (« premiers ») : l'homme, les singes, ainsi que d'autres animaux avec lesquels nous serions sans doute très surpris de nous voir associés.

Les primates sont caractérisés par des mains et parfois des pieds préhensiles, c'est-à-dire que leur pouce peut s'opposer aux autres doigts. L'extrémité de leurs doigts est recouverte d'un ongle plat, et non d'une griffe ou d'un sabot. Le cerveau est développé, et la vue prime sur l'odorat. Un grand nombre d'autres critères anatomiques, moins évidents, les distinguent.

Les primates se divisent en neuf familles. Certains de leurs membres ne présentent que si peu de caractéristiques primatiales qu'on a du mal à les imaginer comme tels, et pourtant primates ils sont. L'une des familles, les tupaïidés, comprend la musaraigne arboricole insectivore ! Il y a également les lémuriens – des animaux nocturnes vivant dans les arbres, au museau pareil à celui du renard, et à l'apparence générale de l'écureuil, que l'on trouve surtout à Madagascar.

Les familles les plus proches de l'homme sont, bien sûr, les singes. Les deux familles de singes d'Amérique, appelés *singes du Nouveau Monde*, sont les cébidés (par exemple, le singe du montreur de foire), et les callithricidés (par exemple, le ouistiti). La troisième, la famille du *Vieux Monde*, est celle des cercopithécidés, dont font partie les différents babouins.

Les grands singes appartiennent tous à une même famille, les pongidés. On les trouve en Afrique et en Asie. Les différences extérieures les plus marquantes qui les distinguent des petits singes sont, bien évidemment, leur plus grande taille mais aussi l'absence de queue. Il en existe quatre types : le gibbon, le plus petit, le plus poilu, aux bras les plus longs, et le plus primitif de toute la famille ; l'orang-outan, plus grand, habite les arbres comme le gibbon ; le gorille, plus grand que l'homme, vit dans les forêts d'Afrique ; et le chimpanzé, également natif d'Afrique, plutôt plus petit que l'homme, est, à part ce dernier, le plus intelligent des primates.

Quant à notre propre famille, celle des hominidés, elle consiste aujourd'hui en un seul genre, et, en fait, en une seule espèce. Linné a appelé cette espèce *Homo sapiens* (« homme sage »), et personne n'a osé changer cette appellation, malgré toutes les objections qu'on peut lui faire.

L'évolution

Il est presque impossible de parcourir la liste des êtres vivants, comme je viens de le faire, sans la quitter intimement convaincu que la vie s'est développée lentement en allant du plus simple vers le plus complexe. On peut classer les embranchements de telle manière que chacun semble ajouter une pierre à l'édifice du précédant. A l'intérieur de chaque phylum, les diverses classes peuvent être arrangées de la même façon ; et à l'intérieur de chaque classe, il en va de même pour les ordres.

De plus, les espèces semblent souvent se confondre, comme si elles étaient encore en train d'évoluer sur des voies proches les unes des autres à partir d'ancêtres communs ayant existé il n'y a pas encore si longtemps. Certaines espèces sont tellement proches l'une de l'autre que, sous certaines conditions, elles sont interfécondes, comme c'est le cas du cheval et de l'âne qui, avec un minimum de bonne volonté, donnent naissance au mulet. On parvient ainsi à accoupler la vache et le bison, le lion et la tigresse. Il existe également des espèces intermédiaires, pour ainsi dire – des créatures qui relient deux plus grands groupes d'animaux entre eux. Le guépard est un chat aux caractéristiques très canines, et la hyène est un chien avec certaines caractéristiques très félines. L'ornithorynque est un mammifère très proche des reptiles. Le *peripatus* semble être moitié ver, moitié mille-pattes. La frontière paraît extrêmement vague lorsque l'on étudie certains animaux à leurs premiers stades de développement. La jeune grenouille ressemble à un poisson ; et il existe un céphalocordé primitif, le *blanoglossus,* découvert en 1825, qui, lorsqu'il est petit, ressemble tellement à un échinoderme qu'on l'a pendant longtemps classé comme tel.

Quand on étudie le développement de l'homme à partir de l'embryon, on s'aperçoit qu'il suit pratiquement le même cheminement que celui qui se produit tout au long des différents embranchements. Cette étude (l'*embryologie*) a vu le jour grâce à Harvey, qui a mis en évidence la circulation du sang. En 1759, le physiologiste allemand Kaspar Friedrich Wolff a pu démontrer que les changements qui interviennent dans l'œuf ne sont en réalité que des stades différents de son développement ; c'est-à-dire qu'à partir de tissus spécialisés, naissent par des changements progressifs des précurseurs non spécialisés, et non pas (comme on le pensait auparavant) par la simple croissance de petites structures déjà spécialisées existant déjà dans l'œuf.

Au cours de son développement, l'œuf n'est tout d'abord qu'une simple cellule (une sorte de protozoaire), puis il devient une petite colonie de cellules (comme l'éponge), chacune d'elles étant capable de se séparer de la colonie et de se développer toute seule : c'est ce qui arrive dans le cas des vrais jumeaux. L'embryon passe alors par un stade à deux couches (comme un cœlentéré), puis reçoit une troisième couche (comme un échinoderme), et continue ainsi de se développer de façon de plus en plus complexe suivant un schéma qui rappelle grosso modo l'évolution des espèces. A un certain stade de son développement, l'embryon humain possède la corde dorsale d'un céphalocordé primitif ; plus tard, il est pourvu de petits sacs branchiaux qui rappellent ceux des poissons, et plus tard encore, il possède la queue et les poils d'un mammifère inférieur.

LES PREMIÈRES THÉORIES

Depuis Aristote, l'idée d'évolution a fleuri chez de nombreux philosophes. Mais la progression du christianisme a mis un frein à ce qui, après tout, n'était encore que spéculation. Le premier chapitre de la Bible affirme péremptoirement que tous les êtres vivants ont été crées « selon leur espèce », ce qui, pris littéralement, signifie que les espèces sont « immuables » et ont aujourd'hui la même forme qu'à l'origine des temps. Même Linné, qui a pourtant été sans doute frappé par la parenté évidente entre les êtres vivants, défendait farouchement le concept de l'immuabilité des espèces.

Le récit littéral de la Création, bien que solidement ancré dans les esprits, a pourtant dû tenir compte des témoignages apportés par les *fossiles* (du latin « creuser »). Dès 1669, le savant danois Nicolaus Steno avait remarqué que les couches profondes (les *strates*) du sol étaient nécessairement plus vieilles que les couches supérieures. En admettant une vitesse raisonnable de formation des roches, il apparut de plus en plus évident que les strates inférieures devaient être *beaucoup* plus âgées que les strates supérieures. On trouvait souvent les restes d'animaux pétrifiés si profondément enfouis qu'ils devaient être beaucoup plus anciens que les quelques millénaires écoulés depuis la Création telle qu'elle était décrite dans la Bible. Les preuves apportées par les fossiles laissaient également supposer que la structure de la Terre avait subi de profonds changements. Dès le VIe siècle av. J.-C., le philosophe grec Xénophane avait noté l'existence de coquillages aux flancs des montagnes et en avait déduit que ces montagnes avaient probablement été submergées dans le passé.

Les défenseurs de l'interprétation littérale de la Bible pouvaient bien sûr maintenir (et ils ne s'en sont d'ailleurs pas privés) que la ressemblance des fossiles avec des organismes ayant existé à une autre époque n'était due qu'au hasard, ou qu'ils avaient été plantés là par le Diable pour tromper les hommes. De telles vues étaient pour le moins peu crédibles, et il leur a donc fallu trouver une hypothèse plus plausible : les fossiles représentaient les restes de créatures noyées pendant le Déluge. Il ne faisait aucun doute que des coquillages retrouvés au sommet des montagnes confirmaient cette théorie, puisqu'il est dit dans la Bible que les eaux ont recouvert toutes les montagnes.

Mais à y regarder de plus près, il fallait bien se rendre à l'évidence : beaucoup de fossiles ne ressemblaient à aucune espèce vivante. John Ray, l'auteur de la première classification, s'est demandé s'il ne s'agissait pas là d'espèces disparues. Le naturaliste suisse Charles Bonnet est allé plus loin et a suggéré, en 1770, que les fossiles étaient effectivement des restes d'espèces disparues, détruites au cours de catastrophes géologiques bien antérieures au Déluge.

C'est un géomètre anglais du nom de William Smith, cependant, qui a jeté les bases scientifiques de l'étude des fossiles et de la préhistoire (la *paléontologie*). Alors qu'il participait à des travaux d'excavation liés à l'aménagement d'un canal en 1791, il fut frappé par le fait que la masse rocheuse qu'il traversait était divisée en strates, et que chacune d'elles contenait ses propres fossiles caractéristiques. Il devenait donc possible de classer ces fossiles par ordre chronologique, en fonction de leur emplacement dans la série des couches successives, et d'associer chaque fossile à un certain type de strate caractéristique d'une période donnée dans l'histoire géologique.

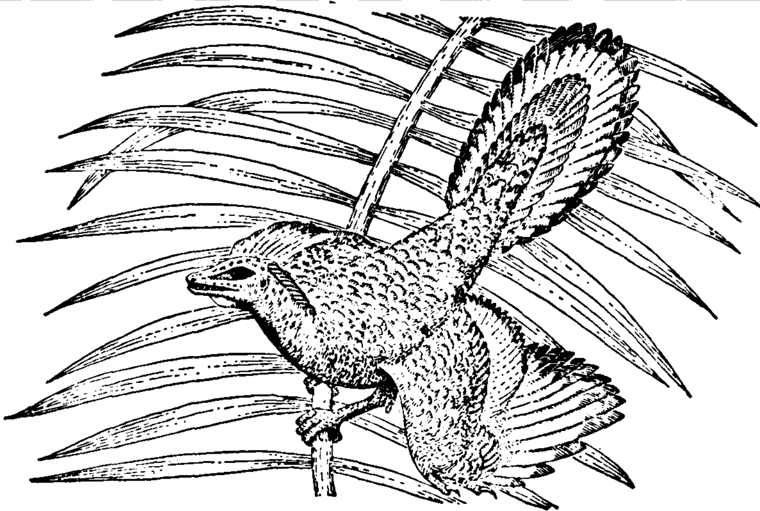
Vers 1800, Cuvier (à qui l'on doit la notion de phylum) a classifié les fossiles en se basant sur le système de Linné, permettant ainsi d'étendre l'étude de l'anatomie comparée aux temps les plus reculés. Bien que de nombreux fossiles correspondent à des espèces ou à des genres disparus, on pouvait facilement les classer à l'intérieur des phylums connus ; ils faisaient donc bien partie intégrante du grand schéma général de la vie telle qu'elle s'était développée sur Terre. En 1801, par exemple, Cuvier a étudié un fossile découvert vingt ans auparavant et qui possédait des doigts très longs ; il a pu démontrer qu'il s'agissait des restes d'un animal volant aux ailes recouvertes d'un cuir épais. Il a également pu montrer à partir de la structure osseuse de ces *ptérodactyles* (« doigts en forme d'ailes »), comme il les a baptisés, qu'ils n'en étaient pas moins des reptiles, appartenant sans aucun doute à la famille des serpents, lézards, alligators et tortues d'aujourd'hui.

D'autre part, plus la strate dans laquelle on trouvait un fossile était profonde, plus ce dernier était vieux, et plus sa structure paraissait primaire et peu développée. Les fossiles représentaient parfois des formes intermédiaires reliant deux groupes qui, à première vue, paraissaient sans rapport. Un exemple particulièrement frappant, découvert après la mort de Cuvier, se trouve être un oiseau primitif appelé *archéoptérix* (du grec « vieille aile »). Cet animal maintenant disparu possédait des ailes et des plumes, et avait une queue pareille à celle d'un lézard, mais dotée de plumes en éventail, et un bec hérissé de dents de reptile ! Il s'agissait clairement d'un animal à mi-chemin entre un reptile et un oiseau (figure 16.3).

Cuvier n'en pensait pas moins que des catastrophes naturelles, et non pas l'évolution, étaient responsables de la disparition de certaines formes de vie ; mais dans les années 1830, les nouveaux concepts de Charles Lyell sur les fossiles et sur l'histoire géologique, exposés dans son livre, *Principes de géologie*, rallièrent la communauté scientifique à son point de vue.

L'élaboration d'une théorie de l'évolution devenait nécessaire pour expliquer toutes les découvertes de la paléontologie.

Figure 16.3. L'archéoptérix.



Si les animaux avaient évolué à partir d'une forme ou d'une autre, quelle pouvait bien en être la cause ? C'est sur cette question qu'achoppaient toutes les tentatives menées jusque-là pour expliquer la diversité du monde vivant. Le premier à proposer un élément de réponse a été le naturaliste français Jean-Baptiste de Lamarck. Il publia en 1809 sa *Philosophie zoologique*, dans laquelle il suggérait que l'environnement provoque chez les organismes de petites modifications qui se transmettent ensuite chez leurs descendants. Pour illustrer son idée, Lamarck prit la girafe (cette dernière venait à peine d'être découverte et faisait sensation à cette époque). Supposons, disait-il, qu'un animal primitif, ressemblant à l'antilope et se nourrissant de feuilles sur les branches des arbres, vienne à manquer de nourriture facilement accessible, et doive allonger son cou le plus possible pour l'atteindre. En prenant l'habitude d'allonger son cou, sa langue, et ses jambes, cet animal étirerait progressivement ces parties de son corps. Il transmettrait alors ces nouvelles caractéristiques à ses descendants, et ainsi de suite. Peu à peu, après des générations et des générations d'un tel allongement, l'antilope primitive se serait transformée en girafe.

Cette notion de l'*héritage des caractères acquis* de Lamarck s'est vite trouvée en butte à mille difficultés. Qu'avait donc fait la girafe pour obtenir son pelage tacheté, par exemple ? Il était évident qu'aucune action, délibérée ou non, de sa part n'avait pu provoquer ce changement. Un expérimentateur sceptique, le biologiste allemand August Friedrich Leopold Weismann, coupa la queue à des souris pendant plusieurs générations, pour en arriver à la conclusion que les souris de la dernière génération possédaient une queue qui n'était pas une fraction de millimètre plus courte que celle de la première génération. En fait, il aurait pu faire l'économie de bien des efforts s'il avait pensé à l'exemple de la circoncision des garçons dans la religion juive, qui, après plus de cent générations, n'a pas provoqué le moindre rétrécissement du prépuce.

Weismann avait déjà observé dès 1883 que les cellules germinales, celles qui vont donner naissance aux spermatozoïdes et aux ovules, se séparaient du reste de l'embryon à un stade très jeune, et restaient relativement peu spécialisées. A partir de cette observation et de ses expériences sur les queues de souris, Weismann a déduit le concept de *continuité du plasme germinal*. Ce dernier (c'est-à-dire le protoplasme qui constitue les cellules germinales) était caractérisé, pensait-il, par une existence indépendante, continue, génération après génération, alors que le reste de l'organisme ne constituait en quelque sorte qu'une demeure temporaire, construite et détruite à chaque génération. Le plasme germinal déterminait les caractéristiques du corps, mais n'était pas lui-même soumis à l'influence du corps. En cela, sa pensée était à l'opposé de celle de Lamarck, tout en étant aussi erronée, bien qu'en gros la vérité soit plus proche des vues de Weismann que de celles de Lamarck.

Bien que rejeté par la plupart des biologistes, le lamarckisme s'est prolongé jusqu'au XXe siècle, et a même connu un certain regain de popularité, temporaire certes, sous la forme du lysenkisme (la modification génétique des plantes provoquée par certains traitements) en Union Soviétique. (Trofim Denisovitch Lysenko, le fondateur de cette doctrine, fut très puissant sous Staline, et conserva beaucoup d'influence sous Khrouchtchev, mais fut éclipsé après le limogeage de ce dernier en 1964.)

Aujourd'hui, les généticiens n'excluent pas que l'environnement puisse engendrer, chez des organismes peu complexes, des changements

transmissibles, mais les idées lamarckiennes en tant que telles furent balayées par la découverte des gènes et des lois de l'hérédité.

LA THÉORIE DE DARWIN

En 1831, Charles Darwin, un jeune gentilhomme anglais un peu dilettante qu'une existence oisive commençait à lasser, recherchait désespérément quelque chose à faire pour combattre l'ennui qui le gagnait ; un capitaine au long cours et un professeur de Cambridge le persuadèrent de les accompagner comme naturaliste au cours d'un voyage autour du monde qui devait durer cinq ans. L'expédition avait pour but principal l'étude des côtes continentales, mais devait être aussi l'occasion d'observer la faune et la flore en cours de route. A l'âge de vingt ans, Darwin allait faire sur le *Beagle* la plus importante croisière dans l'histoire de la science.

Pendant que le navire voguait lentement le long de la côte est de l'Amérique du Sud, puis de la côte ouest, Darwin accumulait méticuleusement une masse stupéfiante d'informations sur les diverses formes de vie végétale et animale qu'il rencontrait. Sa découverte la plus remarquable eut lieu dans un groupe d'îles du Pacifique, situées à un peu plus de mille kilomètres à l'ouest de la république de l'Équateur, appelées les Galapagos, car elles étaient habitées par des tortues géantes (*Galapagos* vient d'un nom espagnol qui signifie « tortue »). C'est une variété de pinsons vivant sur ces îles qui retint le plus l'attention de Darwin pendant le séjour de cinq semaines qu'il y effectua ; aujourd'hui encore, on les connaît sous le nom de *pinsons de Darwin.* Il découvrit que ces oiseaux se divisaient en au moins quatorze espèces différentes qui se distinguaient surtout les unes des autres par la taille et la forme du bec. Ces espèces n'existaient nulle part ailleurs, mais elles semblaient s'apparenter à une espèce habitant le continent sud-américain.

Qu'est-ce qui pouvait bien expliquer les caractéristiques particulières des pinsons vivant sur ces îles ? Pourquoi étaient-ils différents des pinsons ordinaires, et pourquoi étaient-ils divisés en quatorze espèces ? Darwin jugea que l'hypothèse la plus raisonnable était que toutes descendaient de la même espèce de pinson continental, et s'étaient différenciées pendant leur longue période d'isolement sur les îles. La différenciation provenait des diverses méthodes employées par les pinsons pour trouver leur pitance. Trois des espèces sur les Galapagos se nourrissaient encore de graines, comme l'espèce continentale, mais chacune mangeait un type de graine différent, et elles variaient également par la taille : une espèce plutôt grande, une moyenne, et la troisième petite. Deux autres espèces se nourrissaient de cactus ; la plupart des autres vivaient d'insectes.

La modification des habitudes alimentaires chez ces pinsons va hanter Darwin pendant plusieurs années. En 1838, un début de solution va lui être fourni par un livre publié quarante ans auparavant par le pasteur anglais Thomas Robert Malthus. Dans son *Essai sur le principe de population,* Malthus défend l'idée selon laquelle une population s'accroît toujours au point d'excéder ses capacités à se nourrir, et voit donc ses rangs décroître pour cause de famine, de maladies ou de guerre. C'est dans cet ouvrage que Darwin va trouver l'expression « la lutte pour l'existence », que ses théories vont plus tard rendre célèbre. Darwin réalise tout à coup que, appliquée à ses pinsons, cette lutte pour l'existence peut agir comme un

mécanisme favorisant les individus les mieux adaptés. Lorsque les pinsons qui ont colonisé les Galapagos se sont multipliés au point d'excéder par leur nombre les ressources locales en nourriture, les seuls survivants ont été les oiseaux les plus résistants, c'est-à-dire ceux particulièrement bien équipés par la nature pour trouver des graines, ou encore ceux capables de s'adapter à une nouvelle alimentation. Un oiseau possédant des caractéristiques légèrement différentes du pinson moyen, et par exemple capable de manger des graines plus grosses, plus dures ou, mieux encore, des insectes, un tel spécimen trouverait des réserves de nourriture abondantes. Un oiseau muni d'un bec quelque peu plus fin ou plus long pourrait atteindre des aliments inaccessibles pour d'autres. De tels spécimens, ainsi que leurs descendants, verraient leur nombre s'accroître aux dépens de la population originale. Chacun de ces nouveaux types de pinson trouverait un nouveau « créneau » dans l'environnement et s'y installerait. Sur les îles Galapagos, pratiquement dépourvues à l'origine de tout oiseau, ces « créneaux » ne demandaient qu'à être occupés, puisqu'il n'existait aucun concurrent pour barrer la route. Sur le continent sud-américain, en revanche, tous les créneaux étaient pris, et le pinson ancestral avait fort à faire rien que pour subsister. Il n'avait engendré aucune autre espèce.

Darwin a suggéré que chaque génération était constituée d'un ensemble d'individus variant au hasard par rapport à la moyenne. Certains étaient un peu plus grands ; d'autres possédaient des organes de forme un peu différente ; d'autres encore avaient des capacités un peu au-dessus ou un peu au-dessous de la normale. Ces différences étaient peut-être infinitésimales, mais les pinsons que les caractéristiques rendaient un tout petit peu plus à même d'affronter l'environnement tendraient à vivre un tout petit peu plus longtemps, et donc à procréer également plus longtemps. Arrivée à un certain point, une telle accumulation de caractères favorables serait couplée à une incapacité de se reproduire avec le type original ou avec d'autres variantes de ce même type, et c'est ainsi que naîtrait une nouvelle espèce.

Darwin a appelé ce processus la *sélection naturelle*. D'après lui, ce n'est pas en l'allongeant que la girafe s'est dotée d'un long cou, mais parce que certaines girafes naissaient avec un cou plus long que celui de leurs congénères ; plus le cou est long, plus la girafe peut facilement atteindre sa nourriture. La sélection naturelle a permis à l'espèce dotée d'un long cou de gagner la bataille dans la lutte pour l'existence. Ce processus explique aussi aisément le pelage tacheté de la girafe : un animal possédant de telles taches peut se fondre plus facilement dans un paysage baigné de soleil ; il améliore ainsi ses chances d'échapper à la convoitise d'un lion en quête de gibier.

Les thèses de Darwin sur la création des espèces apportaient aussi certains éclaircissements sur la distinction entre espèces ou entre genres. L'évolution des espèces est un processus continu et, bien sûr, très lent. Il existe un nombre indéterminé d'espèces dont les membres sont, en ce moment même, en train de tendre imperceptiblement vers l'établissement de nouvelles espèces.

Darwin a passé plusieurs années à accumuler ses observations et à bâtir sa théorie. Il était persuadé qu'elle ferait trembler les fondements de la biologie et bousculerait les idées de la société sur la place que tient l'homme dans l'ordre des choses, et il tenait sans cesse à s'assurer de la solidité du

terrain sur lequel il avançait. Darwin a commencé à écrire des notes sur le sujet et à y consacrer toute son énergie en 1834, avant même d'avoir lu Malthus ; et il travaillait encore en 1858 sur un livre traitant du même sujet. Ses amis (dont Lyell, le géologue) savaient qu'il y travaillait avec acharnement ; plusieurs d'entre eux avaient lu ses manuscrits. Ils le suppliaient d'activer leur publication, craignant qu'un autre ne le devance. Mais Darwin n'a pas voulu (ou n'a pas pu) faire plus vite, et c'est un autre qui le coiffa au poteau.

Cet homme, c'est Alfred Russel Wallace, quatorze ans plus jeune que Darwin. La vie de Wallace avait suivi un cheminement assez parallèle à celle de Darwin. Lui aussi avait fait partie d'une expédition scientifique autour du monde alors qu'il n'était encore qu'un jeune homme. Dans les îles de la Sonde, il avait observé que les plantes et les animaux des îles situées à l'est différaient fondamentalement de ceux des îles situées à l'ouest. On pouvait tracer une frontière géographique entre les deux formes de vie ; la ligne passait entre Bornéo et les Célèbes, par exemple, et entre les petites îles de Bali et de Lombok plus loin au sud. Cette ligne s'appelle encore aujourd'hui *ligne de Wallace*. (Plus tard dans sa vie, Wallace a divisé le monde en six grandes régions, caractérisées par différentes espèces animales, une division qui a encore cours de nos jours, malgré quelques modifications mineures.)

Il s'avère que les mammifères qui habitent les îles situées à l'est et l'Australie sont beaucoup plus primitifs que ceux qui vivent dans les îles situées à l'ouest et en Asie – ou même dans le reste du monde. Tout laisse à penser que l'Australie et les îles orientales se sont séparées de l'Asie il y a très longtemps, à une époque où seuls existaient des mammifères primitifs, et que les mammifères à placenta se sont développés plus tard seulement en Asie. La Nouvelle-Zélande est probablement restée isolée pendant plus longtemps encore, car on n'y trouve aucun mammifère ; seuls, des oiseaux primitifs (qui d'ailleurs ne volent même pas) l'habitent, dont le survivant le plus connu aujourd'hui est le kiwi.

Comment les mammifères supérieurs étaient-ils apparus en Asie ? Wallace se pencha sur cette question en 1855. En 1858, lui aussi prit connaissance des thèses de Malthus ; et lui aussi en tira les mêmes conclusions que Darwin. Mais à la différence de Darwin, il ne mit pas quatorze ans à les écrire. Dès qu'il eut les idées claires sur le sujet, il s'assit à son bureau et rédigea une note en deux jours. Il décida de faire relire son manuscrit par un biologiste compétent et réputé, et son choix se porta sur Charles Darwin.

Lorsque ce dernier reçut le manuscrit, ce fut exactement comme si la foudre tombait à ses pieds. Il y retrouvait toutes ses idées comme si lui-même les avait écrites. Sans attendre un seul instant, il communiqua l'article de Wallace à d'autres scientifiques de renom, et proposa à Wallace de collaborer à la rédaction de rapports résumant leurs conclusions communes. Ces dernières furent l'objet de communications publiées dans *La Revue de la société linnéenne* en 1858.

Le livre de Darwin parut finalement une année plus tard, sous le titre complet *L'origine des espèces au moyen de la sélection naturelle ou la lutte pour l'existence dans la nature*. Nous le connaissons aujourd'hui sous son titre abrégé, *L'origine des espèces*.

La théorie de l'évolution s'est modifiée et affinée depuis l'époque de Darwin, profitant des connaissances acquises sur les mécanismes de

l'hérédité , des gènes, et des mutations (voir chapitre 13). Ce n'est pas avant 1930, en effet, que le statisticien et généticien anglais Ronald Aylmer Fisher est parvenu à prouver que la génétique mendélienne décrit le mécanisme nécessaire pour expliquer l'évolution par la sélection naturelle. Ce n'est qu'alors que la théorie de l'évolution a conquis ses véritables lettres de noblesse.

Bien évidemment, les progrès accomplis dans d'autres sciences ont permis d'affiner les thèses de Darwin. Les découvertes faites dans le domaine de la tectonique des plaques (voir chapitre 4) ont apporté une meilleure connaissance des forces qui sous-tendent l'évolution et expliquent l'existence d'espèces similaires dans des lieux séparés par de grandes distances. L'analyse détaillée des protéines et des acides nucléiques a permis de suivre l'évolution au niveau moléculaire et de déduire certaines différences entre molécules (comme je l'explique plus loin dans ce chapitre), et certaines différences dans les rapports entre organismes.

Naturellement, dans un domaine aussi complexe que l'évolution et le développement des organismes vivants sur une période couvrant plusieurs milliards d'années, la controverse sur les détails du mécanisme continue à faire rage. C'est ainsi que dans les années soixante-dix, des biologistes comme Stephen Jay Gould ont émis l'hypothèse de l'*évolution ponctuée.* Pour ces chercheurs, l'évolution ne s'inscrit pas dans le cadre d'un développement lent, d'un processus plus ou moins uniforme et continu. Ils pensent plutôt qu'il existe de longues périodes de relative immuabilité, interrompues par des situations dans lesquelles se produisent des changements comparativement soudains et accentués (pas en une nuit, bien sûr, mais au cours de quelques centaines de milliers d'années, ce qui est rapide à l'échelle des processus de l'évolution).

Néanmoins, aucun biologiste de renom ne doute aujourd'hui de la validité du concept de l'évolution. Les bases sur lesquelles se fondent les thèses de Darwin demeurent inébranlées ; ces thèses ont bien évidemment influencé tous les domaines de la science – la physique, la biologie, et les sciences sociales.

OPPOSITION À LA THÉORIE

L'annonce de la théorie de Darwin a naturellement provoqué pas mal de remous. Dans un premier temps, même, bon nombre de scientifiques se sont violemment élevés contre. Elle a ainsi trouvé son plus farouche ennemi en la personne du zoologiste anglais Richard Owen, qui avait succédé à Cuvier comme grand expert mondial des fossiles et de leur classification. Owen a eu recours aux manœuvres les plus basses pour combattre le darwinisme, se cachant sournoisement derrière sa réputation alors qu'il encourageait d'autres à mener la lutte ; il a même écrit des pamphlets anonymes contre la théorie en se citant personnellement comme autorité en la matière.

Le naturaliste anglais Philip Henry Gosse a essayé de se tirer du dilemme en suggérant que la Terre avait été créée par Dieu, y compris les fossiles, pour mettre à l'épreuve la foi des hommes. Bien évidemment, pour quiconque a deux sous de bon sens, suggérer que Dieu s'amuse à jouer de tels tours pendables aux pauvres humains que nous sommes relève du blasphème, et dépasse tout ce que Darwin avait bien pu suggérer.

Ces attaques firent long feu, l'opposition du monde scientifique diminua progressivement, et en l'espace de vingt ans, disparut quasi complètement. Mais les opposants hors de la communauté scientifique continuèrent le combat beaucoup plus longtemps, et beaucoup plus vigoureusement. Les fondamentalistes (c'est-à-dire les défenseurs d'une interprétation littérale de la Bible) étaient scandalisés que l'on puisse insinuer que l'homme descende du singe. Benjamin Disraeli (le futur premier ministre britannique) lança un jour : « La question que notre société doit donc résoudre maintenant est la suivante : l'homme est-il un singe ou bien un ange ? Moi, je suis du côté des anges. » Les hommes d'Église, ralliant évidemment le clan des anges, menèrent le combat contre Darwin.

Darwin n'était pas d'un tempérament à s'impliquer personnellement dans la controverse, mais il trouva un remarquable champion en la personne de l'éminent biologiste Thomas Henry Huxley. Surnommé « le bouledogue de Darwin », Huxley s'est battu vaillamment sur le front des amphithéâtres universitaires anglais. Il remporta sa victoire la plus éclatante en 1860 au cours d'un débat célèbre qui l'opposa à Samuel Wilberforce, un évêque anglican, mathématicien et redoutable orateur surnommé « Sam l'enjôleur ».

Après avoir, semblait-il, conquis l'auditoire, Wilberforce se tourna enfin vers son adversaire, un homme imposant et au sens de l'humour pratiquement inexistant. Le compte rendu du débat rapporte que Wilberforce « pria (Huxley) de bien vouloir indiquer aux personnes présentes si c'était par son grand-père ou par sa grand-mère qu'il descendait du singe ».

Alors que l'auditoire riait aux éclats, Huxley se tourna vers son voisin et lui souffla : « Le Tout-Puissant vient de me le livrer pieds et poings liés » ; il se leva lentement et répondit : « Puisque l'on me pose la question de savoir si je préférerais avoir un misérable petit singe comme grand-père, ou bien un homme comblé par la nature, jouissant de grands privilèges et d'une grande influence, et qui pourtant n'hésite pas à utiliser ces privilèges et cette influence dans le seul but d'introduire le ridicule dans une discussion scientifique sérieuse, j'affirme sans aucune hésitation que ma préférence va vers le singe. »

Apparemment, la répartie cinglante de Huxley brisa non seulement l'élan de l'évêque Wilberforce, mais mit également les fondamentalistes sur la défensive. En fait, la victoire du point de vue des darwiniens était tellement évidente que, lorsque Darwin mourut en 1882, il fut enterré à l'abbaye de Westminster, où sont enterrés la plupart des grands hommes anglais, entouré d'un respect quasi général. C'est aussi en son honneur que fut baptisée la ville de Darwin, dans le Nord de l'Australie.

Un autre défenseur ardent des thèses évolutionnistes fut le philosophe anglais Herbert Spencer, qui popularisa l'expression « la survie du mieux adapté » et le mot « évolution », un terme que Darwin lui-même utilisait peu. Spencer a tenté d'appliquer la théorie de l'évolution au développement des sociétés humaines (on le considère d'ailleurs comme le fondateur de la sociologie). Néanmoins, ses arguments n'étaient pas valides car les modifications biologiques impliquées dans l'évolution sont sans rapport avec les changements qui se produisent dans les sociétés ; de plus, c'est contre sa volonté que certaines de ses thèses ont été invoquées à mauvais escient pour soutenir la guerre et le racisme.

Aux États-Unis, un conflit dramatique éclata en 1925 au sujet de l'évolution ; les anti-évolutionnistes gagnèrent la bataille, mais perdirent la guerre.

La législature de l'État du Tennessee avait voté une loi interdisant l'enseignement de la théorie de l'évolution dans les écoles publiques. Afin d'en tester la constitutionnalité, des chercheurs et des enseignants persuadèrent un jeune professeur de biologie, John Thomas Scopes, d'enseigner le darwinisme à ses élèves. Scopes fut immédiatement accusé d'avoir transgressé la loi, et son procès eut lieu à Dayton, une petite ville du Tennessee où se trouvait son lycée. Le procès eut un retentissement mondial.

La population locale et le juge étaient farouchement opposés aux thèses de l'évolution. William Jennings Bryan, avocat célèbre, trois fois candidat à la Maison Blanche, et fondamentaliste convaincu, siégeait dans le fauteuil de l'avocat général. Scopes fut défendu par le célèbre avocat Clarence Darrow, entouré d'une pléïade de confrères.

En gros, le procès fut décevant, car le juge refusa à la défense le droit de faire venir à la barre des hommes de science pour expliquer la théorie de Darwin, et n'admit que les dépositions ayant trait directement au chef d'inculpation : oui ou non, Scopes avait-il discuté l'évolution en classe ? Mais le problème de fond fut néanmoins soulevé au cours d'une séance du procès lorsque Bryan, contre l'avis de ses confrères de l'accusation, accepta de se soumettre à un contre-interrogatoire pour présenter la position fondamentaliste. Darrow n'eut aucun mal à démontrer que Bryan ignorait tout des derniers développements scientifiques de son temps, et n'avait de la religion en général et de la Bible en particulier que quelques notions rudimentaires qui lui restaient du catéchisme.

Scopes fut déclaré coupable, et condamné à payer cent dollars d'amende. (Le verdict fut annulé plus tard pour vice de forme par la cour suprême de l'État du Tennessee.) Mais les fondamentalistes (et l'État du Tennessee) s'étaient tellement ridiculisés aux yeux du monde que les anti-darwiniens furent contraints de mener un combat d'arrière-garde, et restèrent dans l'ombre pendant près d'un demi-siècle.

Mais les puissances de l'ombre et de l'ignorance ne sont jamais totalement vaincues ; et dans les années soixante-dix, les anti-darwiniens ont fait un retour en force et mènent un combat encore plus insidieux contre les vues de la science sur l'Univers. Ils n'invoquent plus à présent la nécessité d'interpréter littéralement la Bible (cette position avait au moins pour elle sa franchise, bien qu'elle soit totalement discréditée), mais prétendent à la respectabilité scientifique. Ils font de vagues références à un « Créateur » et se gardent bien d'utiliser les formules bibliques. Ils soutiennent depuis peu que la théorie de l'évolution est entachée d'erreurs et ne peut donc décrire la réalité et que, par conséquent, le *créationnisme* seul est dans le vrai.

Ils n'hésitent pas, pour démontrer les prétendues erreurs de la théorie darwinienne, à en citer certains passages hors de contexte, et à leur donner des interprétations erronées, transgressant par là même les règles bibliques qui proscrivent les faux témoignages. Et s'ils proclament maintenant que leur vue est la seule vraie, c'est sans avoir jamais produit le moindre élément de preuve en faveur de leur cause, qu'ils appellent, assez pompeusement, le « créationnisme scientifique ».

C'est sans vergogne qu'ils demandent un « droit de réponse » dans les écoles, et exigent que tout enseignant ou tout livre présentant la théorie de

l'évolution présente également les thèses du « créationnisme scientifique ». A l'heure où j'écris ces lignes, ils n'ont encore gagné aucune bataille devant les tribunaux des États-Unis, mais leurs porte-parole, soutenus par des croyants sincères mais non scientifiques, et pour qui tout chose hors de la Bible n'est qu'un épais brouillard qu'il est préférable de ne pas percer, menacent les établissements scolaires, les bibliothèques et les législateurs pour les forcer, purement et simplement, à censurer, voire éliminer la science.

Les résultats risquent d'être désastreux, car les idées créationnistes selon lesquelles la Terre n'a que quelques milliers d'années, comme d'ailleurs l'Univers tout entier – la vie serait apparue soudain avec toutes ses espèces distinctes dès le commencement du monde – sont autant d'insultes à l'astronomie, à la physique, la géologie, la chimie et la biologie, et donneraient lieu à une génération de jeunes gens et de jeunes filles à l'esprit noyé dans les ténèbres d'une nuit peut-être sans lendemain.

LES ARGUMENTS DE LA THÉORIE

L'un des arguments des créationnistes tient au fait que personne n'a jamais pu observer les forces de l'évolution à l'œuvre. Il semblerait à première vue qu'ils tiennent là un argument des plus irréfutables, et pourtant, là encore, ils se trompent.

De fait, s'il fallait absolument démontrer le darwinisme d'une manière ou d'une autre, ce ne sont pas les exemples de sélection naturelle qui manquent ; et ces exemples se sont déroulés devant nos propres yeux (nous savons maintenant ce qu'il faut regarder). C'est la terre natale de Darwin qui en a fourni la première preuve éclatante.

Il s'avère qu'en Angleterre, il existe deux variétés de géomètres du bouleau (des lépidoptères), une claire et une foncée. A l'époque de Darwin, la variété claire prédominait parce qu'elle se confondait plus facilement avec l'écorce des arbres recouverte de lichens, également de couleur claire. Cette *coloration protectrice* la protégeait plus souvent contre les prédateurs que l'autre espèce plus foncée, et donc plus facilement visible. Mais l'Angleterre s'est progressivement industrialisée, la pollution a tué les lichens, et l'écorce des arbres est devenue plus sombre. C'était alors l'espèce foncée qui devenait moins visible et se trouvait donc mieux protégée. Dès lors, cette dernière a prédominé – par le biais de la sélection naturelle.

En 1952, le parlement britannique a promulgué une série de lois visant à combattre la pollution et à protéger l'environnement. Le niveau de la pollution a baissé, le lichen clair a pu de nouveau pousser sur l'écorce des arbres, et la population de lépidoptères clairs a de nouveau augmenté. La théorie de l'évolution explique parfaitement ces événements ; mais une bonne théorie se doit de pouvoir non seulement expliquer le présent, mais également prévoir le futur.

Le cours de l'évolution

En se basant sur l'étude des fossiles, les paléontologues ont divisé l'histoire de la Terre en plusieurs *ères*. Ce sont des géologues anglais du

XIXe siècle, dont Lyell, Adam Sedgwick et Roderick Impey Murchison, qui ont effectué ce travail et qui ont donné un nom aux ères successives. Elles débutent il y a quelque six cents millions d'années avec les premiers fossiles sur la nature desquels on ne puisse se méprendre (c'est-à-dire datant d'une époque où tous les phylums existaient déjà sur Terre, sauf les cordés). Ces premiers fossiles ne sont évidemment pas les premières formes de vie. Bien souvent en effet, ce ne sont que les parties dures d'un organisme qui se fossilisent, et l'on ne peut donc retrouver que des fossiles d'animaux dotés d'une coquille ou d'une ossature. On s'aperçoit que même les plus simples et les plus vieilles de ces créatures sont déjà à un stade d'évolution avancé. Une preuve en a été apportée en 1965, lorsqu'ont été découverts les restes fossilisés de petits coquillages proches de la palourde ; ils dataient d'il y a environ 720 millions d'années.

Les paléontologues sont aujourd'hui équipés pour aller encore plus loin. Il est tout à fait raisonnable, en effet, de penser que la vie unicellulaire a existé bien avant l'apparition d'animaux à coquille ; on a d'ailleurs retrouvé des traces d'algues bleu-vert et de bactéries dans des sédiments âgés d'un milliard d'années et plus. En 1965, le paléontologue américain Elso Sterrenberg Barghoorn a découvert de minuscules objets ressemblant à des bactéries (des *microfossiles*) dans des roches datant de plus de trois milliards d'années. Elles sont si petites qu'on doit étudier leur structure au microscope électronique.

Lorsqu'on voyage dans le temps jusque vers les origines de la vie, il semblerait que l'évolution chimique ait commencé dès que la Terre a pris sa forme actuelle, c'est-à-dire il y a quelque 4,6 milliards d'années. Un milliard d'années plus tard, cette évolution chimique avait atteint un stade où des systèmes assez complexes pour qu'on les appelle vivants se sont formés. A cette époque, l'atmosphère de la Terre en était encore à se réduire et ne contenait aucune trace significative d'oxygène (voir chapitre 5). Les formes de vie les plus reculées devaient donc être adaptées à cet environnement, et leurs descendants ont survécu jusqu'à nos jours.

En 1970, Carl R. Woese a entrepris l'étude détaillée de certaines bactéries qui n'existent qu'en l'absence totale d'oxygène. Certaines d'entre elles réduisent l'oxyde de carbone en méthane ; on les appelle donc des *méthanogènes* (« producteurs de méthane »). Certaines bactéries sont le siège d'autres réactions qui produisent l'énergie nécessaire à la vie sans recours à l'oxygène. Woese les réunit sous le terme d'*archéobactéries* (« vieilles bactéries ») et suggère qu'on les classe dans un règne à part.

Mais après que la vie se fut établie sur Terre, la nature de l'atmosphère a changé – au début, très lentement. Il y a environ deux milliards et demi d'années, les algues bleu-vert existaient peut-être déjà, et le processus de la photosynthèse allait provoquer la transition entre une atmosphère azote-oxyde de carbone et une atmosphère azote-oxygène. Il est probable que les eucaryotes existaient voilà environ un milliard d'années, et la vie unicellulaire devait être très diversifiée dans les mers, dominée par des protozoaires aux caractéristiques manifestement animales, qui auraient alors été les formes de vie les plus complexes sur notre globe – eh oui, les monarques du monde.

Pendant les deux milliards d'années qui ont suivi l'apparition des algues bleu-vert, la quantité d'oxygène a probablement augmenté, tout d'abord très lentement. Alors que ce dernier milliard d'années de l'histoire de la Terre commençait à peine, la concentration d'oxygène représentait

peut-être 1 ou 2 % de l'atmosphère, assez pour fournir une source riche en énergie à des cellules animales n'ayant jamais existé auparavant. L'évolution prenait le chemin d'une plus grande complexité ; et il y a donc six cents millions d'années, certains organismes pluricellulaires pouvaient entamer le long processus qui, aujourd'hui, nous a laissé les fossiles.

Les roches sédimentaires les plus anciennes contenant de tels fossiles appartiennent à l'*âge cambrien ;* jusqu'à une époque récente, on reléguait encore les quatre milliards d'années précédentes de l'histoire de notre Terre sous le terme de *Précambrien.* Maintenant qu'on y a trouvé des traces incontestables de vie, on lui a donné le nom plus approprié d'*éon cryptozoïque* (du grec « vie cachée »), alors que les six cents derniers millions d'années forment l'*éon phanérozoïque* (« vie visible »).

L'éon cryptozoïque est lui-même divisé en deux parties : d'abord l'*ère archéozoïque* (« vie ancienne »), dans laquelle on trouve les premières traces de vie unicellulaire ; puis l'*ère protérozoïque* (« jeune vie »).

La limite entre l'éon cryptozoïque et l'éon phanérozoïque est extrêmement marquée. A un certain point dans le temps, on ne trouve pour ainsi dire aucun fossile de taille macroscopique ; et puis soudain, on trouve des organismes complexes appartenant à une douzaine de types différents. Une division aussi nette s'appelle une *non-conformité,* et un tel phénomène fait immanquablement penser à d'éventuelles catastrophes naturelles. Les fossiles auraient dû apparaître plus graduellement, et il est probable que des événements géologiques extrêmement violents en ont effacé toute trace antérieure.

LES ÈRES ET LES ÂGES

Les grandes divisions de l'éon phanérozoïque sont l'*ère paléozoïque* (« vie ancienne »), *le Mésozoïque* (« vie moyenne ») et *le Cénozoïque* (« nouvelle vie »). Les nouvelles méthodes de datation géologique permettent d'affirmer que l'ère paléozoïque a duré environ 350 millions d'années ; le Mésozoïque, 150 millions d'années ; et le Cénozoïque représente les 50 derniers millions d'années de l'histoire de la Terre.

Chaque ère est elle-même divisée en âges. Le Paléozoïque commence, comme je viens de l'indiquer, par l'âge cambrien (d'un nom de lieu au pays de Galles – en fait d'une tribu qui habitait cette région – où les premières strates correspondant à cet âge ont été découvertes). A cette époque, les formes de vie les plus développées sont les coquillages. C'est l'ère des *trilobites,* des arthropodes primitifs dont le crabe fer à cheval est le représentant le plus proche aujourd'hui. Comme cet animal a survécu jusqu'à nos jours sans pratiquement de modifications depuis 200 millions d'années, il est un exemple de ce que l'on appelle parfois un peu exagérément un *fossile vivant.*

Puis vient l'*Ordovicien* (du nom d'une autre tribu galloise), situé il y a 450 à 500 millions d'années, quand les cordés commencent à apparaître sous la forme de *graptolites,* de petits animaux maintenant disparus qui vivaient en colonies. Il est possible qu'ils soient liés au balanoglosse qui, comme les graptolites, appartient aux *hémicordés,* le sous-embranchement le plus primitif dans le phylum des cordés.

On arrive ensuite au *Silurien* (encore une tribu galloise) et au *Dévonien* (du comté anglais, le Devonshire). L'âge dévonien, qui se situe il y a 350

à 400 millions d'années, voit les poissons prendre une place dominante dans les océans, place qu'ils détiennent encore aujourd'hui. C'est également à cette époque que des formes de vie ont commencé à coloniser la terre ferme. Il nous est difficile d'imaginer, et c'est pourtant la stricte vérité, que pendant les cinq sixièmes de son histoire, la vie s'est confinée à l'élément aquatique, et que la terre ferme n'était que vide et désolation. Lorsqu'on pense aux difficultés que représentent le manque d'eau, les températures extrêmes, la force brutale de la pesanteur non modérée par le milieu aquatique, on comprend mieux le fait que le développement sur terre de formes ayant d'abord évolué dans les mers constitue la victoire la plus extraordinaire jamais remportée par la vie sur l'environnement.

L'exode vers les terres a probablement débuté lorsque la compétition pour la nourriture dans les eaux surpeuplées a conduit certains organismes vers les eaux peu profondes, jusqu'alors inoccupées à cause des marées. Comme de plus en plus d'espèces s'habituaient à vivre dans ces régions, la seule solution pour certaines d'entre elles a été de se rapprocher progressivement des côtes, jusqu'à ce que certains organismes mutants deviennent finalement capables de s'établir sur la terre ferme.

Les premières formes de vie à avoir effectué la transition avec succès ont été certaines plantes, voici quelque 400 millions d'années. Les pionniers appartenaient à un groupe maintenant disparu, les *psilopsides* – les premières plantes pluricellulaires. Elles tirent leur nom du mot grec qui signifie « nu » parce que leurs tiges étaient dépourvues de feuilles, une preuve de la nature primitive de ces plantes. Puis des végétaux plus complexes ont réussi à se développer, à tel point qu'il y a 350 millions d'années, les terres étaient recouvertes de forêts. A partir du moment où les végétaux ont pu commencer à pousser sur la terre ferme, les animaux ont pu s'installer à leur tour. En l'espace de quelques millions d'années, les arthropodes, les mollusques et les vers occupaient une bonne partie des terres. Tous ces premiers animaux terrestres étaient petits, car des animaux plus volumineux, sans squelette interne, se seraient effondrés sous le poids de la pesanteur. (Dans la mer, bien sûr, la portance de l'eau annule en grande partie la pesanteur, qui ne joue donc quasiment aucun rôle ; aujourd'hui encore, les plus gros animaux de notre globe vivent dans l'eau.) Les premiers animaux terrestres à conquérir une réelle autonomie de mouvement ont été les insectes ; grâce au développement de leurs ailes, ils ont pu neutraliser l'effet de la pesanteur qui confinait les autres animaux à une lente reptation.

Finalement, 100 millions d'années après la première invasion des terres, une deuxième vague d'animaux est apparue, qui pouvaient se permettre d'être assez massifs grâce à leur ossature interne. Ces nouveaux colonisateurs venus de la mer étaient des poissons appartenant à la sous-classe crossopterygii (« nageoires à franges »). Certains de leurs cousins ont émigré vers les solitudes des grandes profondeurs, comme le cœlacanthe, dont les biologistes, à leur grand étonnement, ont découvert en 1938 des spécimens vivants.

L'invasion des terres par les poissons est le résultat de l'intense compétition qui s'est déroulée chez ces derniers pour subvenir à leurs besoins en oxygène dans les étendues d'eau saumâtre. Grâce aux quantités illimitées d'oxygène présentes dans l'atmosphère, ce sont ceux capables d'en profiter qui ont le mieux survécu lorsque le niveau d'oxygène dans l'eau est descendu au-dessous du seuil critique pour leur survie. Les organes

capables d'emmagasiner l'air ambiant ont alors pris une importance vitale, et les poissons se sont dotés de poches à l'intérieur de leur appareil digestif, dans lesquelles l'air qu'ils avalaient pouvait être stocké. Ces poches se sont muées, dans certains cas, en des poumons rudimentaires. Les dipnoïques (*poissons à poumons*) sont parmi les descendants de ces poissons primitifs, dont quelques espèces existent encore en Afrique et en Australie. Ils vivent dans des eaux stagnantes où d'autres poissons suffoqueraient, et ils peuvent même survivre à des périodes de sécheresse pendant l'été, au cours desquelles leur habitat devient totalement sec. Même les poissons vivant dans les mers, où l'approvisionnement en oxygène ne pose pourtant aucun problème, portent les marques d'une descendance de ces anciennes créatures, car ils sont encore dotés de poches d'air qu'ils utilisent, non pour respirer, mais comme organes de flottaison.

Mais certains de ces poissons à poumons ont poussé l'expérience plus loin, en vivant pendant des périodes plus ou moins longues complètement hors de l'eau. Les espèces crossoptérygiennes qui possédaient les nageoires les plus solides ont le mieux réussi car, sans l'aide du milieu aquatique, il leur a fallu les utiliser pour vaincre la pesanteur. Dès la fin de l'âge dévonien, quelques-uns de ces crossoptérygiens aux rudiments de poumons se sont ainsi retrouvés sur la terre ferme, dressés maladroitement sur quatre pattes trapues, mais encore chancelantes.

Au Dévonien a succédé le *Carbonifère* (« l'âge du charbon »), ainsi nommé par Lyell car c'est la période où d'immenses forêts marécageuses ont envahi la Terre, il y a quelque trois cents millions d'années (la période de végétation la plus intense qu'elle ait sans doute jamais connue) ; finalement, elles ont été enterrées pour devenir les abondants gisements de charbon de notre planète. C'est l'âge des amphibiens ; dès lors, les crossoptériens vivent toute leur vie adulte sur la terre ferme. Puis vient le *Permien* (ainsi nommé d'après une région de l'Oural que Murchison étudia, au terme d'un long voyage à partir de l'Angleterre). Les premiers reptiles font leur apparition. Ils annoncent l'ère *mésozoïque,* pendant laquelle les reptiles dominent tellement la Terre qu'on l'appelle aussi l'*âge des reptiles.*

Le Mésozoïque est divisé en trois âges : le *Trias* (on l'a découvert dans trois strates), le *Jurassique* (du nom des montagnes françaises du Jura), et le *Crétacé* (« calcaire »). C'est pendant le Trias que sont apparus les *dinosaures* (du grec « lézard terrifiant »). Les dinosaures ont connu leur apogée pendant le Crétacé, lorsque *Tyrannosaurus rex* faisait régner la terreur sur son chemin – le plus grand animal carnivore terrestre ayant jamais existé dans l'histoire de notre planète.

C'est pendant le Jurassique que les premiers mammifères et les premiers oiseaux se sont développés, chacun à partir d'un groupe de reptiles différent. Pendant des millions d'années, ces créatures sont demeurées très discrètes et ne connurent qu'un développement limité. Mais la fin du Crétacé a vu la disparition de tous les dinosaures en un temps relativement court, ainsi que celle d'autres grands reptiles que l'on ne classe pas parmi les dinosaures – les ichtyosaures, les plésiosaures et les ptérosaures. (Les deux premiers étaient des reptiles aquatiques ; le troisième, un reptile volant.) De plus, quelques groupes d'invertébrés, tels les ammonites (apparentés au nautile chambré qui vit encore aujourd'hui), ont disparu – comme beaucoup d'autres organismes plus petits, dont plusieurs types d'organismes microscopiques marins.

Certaines estimations indiquent que près de 75 % de toutes les espèces vivant à cette époque ont disparu au cours de ce qui est parfois appelé *la grande mort* à la fin du Crétacé. Comme il semble qu'un immense carnage ait eu lieu parmi les 25 % restants, il ne serait pas surprenant que 95 % de tous les organismes vivant à cette époque aient disparu. Quelque chose s'est produit qui a entraîné la quasi-stérilisation de la planète – mais quoi ?

En 1979, le paléontologue américain Walter Alvarez dirigeait une équipe chargée d'étudier certaines vitesses de sédimentation en mesurant la concentration de divers métaux dans un échantillon prélevé dans des roches sédimentaires d'Italie centrale. L'un des métaux étudiés par la technique de l'activation neutronique se trouvait être l'iridium ; quelle ne fut pas la surprise d'Alvarez de constater que la concentration d'iridium dans une bande mince de l'échantillon était vingt-cinq fois plus élevée que celle des couches immédiatement au-dessus et au-dessous.

D'où l'iridium pouvait-il bien provenir ? Se pouvait-il que la vitesse de sédimentation ait été particulièrement élevée à cet endroit ? Ou bien se pouvait-il qu'il provienne d'une source particulièrement riche en iridium ? Les météores sont plus riches en iridium et en certains autres métaux que la croûte terrestre, et cette section de l'échantillon était également riche en d'autres métaux. Alvarez émit l'hypothèse qu'un météore s'était écrasé là, mais il n'y avait aucun signe de cratère nulle part dans la région.

Les recherches effectuées plus tard ont cependant montré qu'on retrouvait cette même couche riche en iridium dans des endroits distants de plusieurs milliers de kilomètres, et toujours dans des roches datant de la même période. On a pensé qu'un énorme météore s'était peut-être écrasé sur la Terre, que de vastes quantités de matière avaient été disséminées par l'impact dans la haute atmosphère (y compris les débris du météore qui s'était littéralement pulvérisé), et qu'elles étaient lentement retombées, recouvrant ainsi toute la surface du globe.

A quelle époque cela s'est-il produit ? Les roches, dont la matière riche en iridium a été extraite, dataient d'il y a soixante-cinq millions d'années – précisément la fin du Crétacé. De nombreux géologues et paléontologues (mais pas tous, loin de là) ont commencé à accorder un certain crédit à l'hypothèse selon laquelle les dinosaures et les autres organismes qui semblent avoir péri soudainement, pendant la *grande mort* à la fin du Crétacé, ont disparu lors de l'impact catastrophique sur la Terre d'un objet d'environ dix kilomètres de diamètre – soit un petit astéroïde, soit le noyau d'une comète.

Il est possible que de telles collisions aient eu lieu périodiquement, chacune d'elles entraînant une *grande mort.* Celle de la fin du Crétacé n'est que la plus spectaculaire parmi les plus récentes et donc la plus susceptible d'être étudiée en détail. Et bien sûr, il est également possible que des événements similaires se produisent dans le futur, à moins que le développement de nos techniques spatiales ne permette un jour la destruction d'objets menaçants avant qu'ils n'atteignent notre planète. De fait, il semblerait maintenant que des grandes morts se soient produites régulièrement tous les vingt-huit millions d'années. En 1984, l'hypothèse a été émise que le Soleil possède une petite étoile compagnon et que lorsque celle-ci s'approche de son périhélie, tous les vingt-huit millions d'années, elle disloque le nuage de comètes de Oort (voir chapitre 3) et en projette plusieurs millions à l'intérieur du système solaire. On peut très bien imaginer que quelques-unes d'entre elles entrent en collision avec la Terre.

Un tel impact dévaste immédiatement les zones directement touchées, mais l'effet à l'échelle de la planète est dû principalement à l'énorme quantité de poussières soulevées jusque dans la stratosphère – des poussières qui provoquent une longue nuit glacée sur toute la surface du globe et entraînent une fin temporaire de la photosynthèse.

En 1983, l'astronome Carl Sagan et le biologiste Paul Ehrlich ont signalé que, dans l'hypothèse d'une guerre nucléaire, l'explosion de 10 % de l'arsenal actuel d'armes atomiques soulèverait assez de matière dans la stratosphère pour produire artificiellement une telle nuit hivernale, qui pourrait durer assez longtemps pour compromettre toute vie humaine sur Terre – une autre grande mort dont nous pouvons très certainement nous passer.

Quoi qu'il en soit, la mort des grands reptiles qui dominaient la Terre à la fin du Crétacé, quelle qu'en soit la cause, a permis à l'ère cénozoïque qui lui a succédé de devenir *l'âge des mammifères*. Elle est à l'origine du monde tel que nous le connaissons.

LES MODIFICATIONS BIOCHIMIQUES

L'unité des organismes vivants d'aujourd'hui est démontrée en partie par le fait que tous sont composés de protéines fabriquées à partir des mêmes acides aminés. Et une unité du même ordre avec le passé a également été démontrée récemment. La nouvelle science de la *paléobiochimie* (la biochimie des formes de vie primitives) a vu le jour à la fin des années cinquante, lorsqu'on a pu démontrer que certains fossiles datant de trois cents millions d'années contenaient des restes de protéines fabriquées avec précisément les mêmes acides aminés que ceux d'aujourd'hui – glycine, alanine, valine, leucine, acide glutamique, et acide aspartique. Pas un seul des acides aminés « primitifs » ne différait des acides aminés actuels. De plus, on a retrouvé des traces d'hydrate de carbone, de cellulose, de graisses, de porphyrines, et de nouveau rien ne pouvait les distinguer des mêmes substances qui existent aujourd'hui.

Nos connaissances actuelles en biochimie nous permettent de déduire certaines modifications biochimiques qui ont pu jouer un rôle dans l'évolution des animaux.

Prenons l'exemple de l'excrétion des déchets azotés. Apparemment, la façon la plus simple de se débarrasser de l'azote est de l'excréter sous la forme de la petite molécule d'ammoniac (NH_3), qui peut facilement traverser la paroi cellulaire et se retrouver dans le sang. Or, il s'avère que l'ammoniac est extrêmement toxique ; si sa concentration dans le sang excède une partie par million, l'organisme meurt. Pour un organisme marin, cela ne pose pas de problème ; il peut constamment libérer cet ammoniac dans l'Océan par ses branchies. Mais pour un animal terrestre, il est hors de question d'excréter l'ammoniac. S'en débarrasser aussi vite qu'il se forme exigerait une telle quantité d'urine que l'animal se déshydraterait rapidement et mourrait. Un animal terrestre doit donc produire ses déchets azotés sous une forme moins toxique que l'ammoniac. La réponse est l'urée. Cette substance peut être transportée par le sang à des concentrations allant jusqu'à une partie pour mille sans présenter de danger sérieux.

Les poissons éliminent donc les déchets azotés sous forme d'ammoniac, de même que les têtards. Mais lorsqu'un têtard devient une grenouille,

il élimine ces déchets sous forme d'urée. Cette modification dans la chimie de l'organisme est tout aussi cruciale pour la transformation d'une vie aquatique en une vie terrestre que l'est la transformation visible des branchies en poumons.

Une telle modification biochimique s'est probablement produite lorsque les crossoptérygiens ont envahi la terre ferme pour devenir amphibiens. Il y a donc toutes les raisons de penser que l'évolution *biochimique* a joué un rôle aussi important dans le développement des organismes que l'évolution *morphologique* (c'est-à-dire la modification des formes et des structures).

Une autre modification biochimique a été nécessaire pour que le grand pas de l'amphibien au reptile puisse se produire. Si l'embryon d'un œuf de reptile excrétait de l'urée, cette dernière s'accumulerait et atteindrait vite des concentrations toxiques dans la quantité d'eau limitée contenue dans l'œuf. Une modification a pu résoudre ce problème : la formation d'acide urique au lieu d'urée. L'acide urique (une molécule de purine qui ressemble à l'adénine et à la guanine, deux substances impliquées dans la composition des acides nucléiques) est insoluble dans l'eau ; il est donc précipité sous la forme de petits granules qui ne peuvent ainsi pénétrer les cellules.

A l'âge adulte, les reptiles continuent d'éliminer leurs déchets azotés sous forme d'acide urique. Leur urine n'est pas liquide, mais se présente sous la forme d'une matière pâteuse excrétée par un orifice qui s'appelle le cloaque (du mot latin signifiant « égout »).

Les oiseaux et les mammifères qui pondent, tout comme les reptiles, ont conservé le mécanisme d'élimination de l'acide urique et le cloaque. De fait, on appelle souvent les mammifères qui pondent des *monotrèmes* (du grec signifiant « trou unique »).

Les mammifères placentaires, quant à eux, peuvent facilement se débarrasser des déchets azotés de l'embryon, car ce dernier est connecté indirectement au système circulatoire de la mère. L'urée ne pose donc aucun problème aux embryons de mammifères puisqu'elle passe dans le sang de la mère, qui l'élimine par ses reins.

Un mammifère adulte doit excréter des quantités abondantes d'urine pour se débarrasser de son urée. C'est pour cela qu'il est doté de deux orifices : un anus pour éliminer les déchets solides de sa nourriture, et un urètre pour éliminer l'urine.

Ce bref exposé sur le processus d'élimination des déchets azotés démontre que malgré l'unité fondamentale de la vie, il existe des variations systématiques mineures d'une espèce à l'autre. De plus, l'amplitude de ces variations semble augmenter en fonction de l'éloignement, au point de vue de l'évolution, des deux espèces que l'on compare.

Prenons le cas des anticorps. Leur production augmente dans le sang animal lorsque celui-ci réagit à la présence d'une ou plusieurs protéines étrangères, par exemple, celles du sang humain. De tels *antisérums,* lorsqu'ils sont isolés, réagissent fortement au contact du sang humain, le coagulant, mais ne réagissent pas de cette manière en présence du sang d'autres espèces. (Ce phénomène est à la base des tests mis au point pour identifier l'origine humaine du sang, tests utilisés parfois dans le cadre de certaines enquêtes criminelles.) Il est intéressant de noter que les antisérums qui réagissent au sang humain ont une faible réponse vis-à-vis du sang de chimpanzé, alors que ceux qui réagissent fortement au sang

de poulet ont une faible réponse vis-à-vis du sang de canard, etc. La spécificité des anticorps peut ainsi être utilisée pour montrer l'étroitesse des liens entre diverses formes de vie.

De tels tests mettent en évidence, et ceci n'est pas surprenant, des différences mineures dans la molécule très complexe d'une protéine – des différences assez petites chez les espèces étroitement apparentées pour permettre un certain chevauchement dans les réactions des antisérums.

Quand, dans les années cinquante, les biochimistes ont mis au point des techniques permettant de déterminer avec précision la structure des acides aminés formant les protéines, cette méthode de classification des espèces en fonction de la structure des protéines a fait un énorme bond en avant.

En 1965, des études encore plus détaillées ont été faites sur les molécules d'hémoglobine chez divers types de primates, dont l'homme. Des deux types de chaînes peptidiques qui constituent l'hémoglobine, l'une, la *chaîne alpha,* variait peu d'un primate à l'autre ; l'autre, la *chaîne bêta,* présentait des variations considérables. Entre certains primates et l'homme, il n'y avait que six différences dans les acides aminés de la chaîne alpha, mais vingt-trois dans ceux de la chaîne bêta. A en juger par les différences entre les molécules d'hémoglobine, on pense actuellement que l'espèce humaine a divergé des singes il y a environ soixante-quinze millions d'années, c'est-à-dire à l'époque où les ancêtres des chevaux et des ânes ont eux-mêmes divergé.

Des distinctions encore plus nettes peuvent être établies en comparant les molécules de *cytochrome c,* une molécule de protéine riche en fer comportant environ cent cinq acides aminés, que l'on trouve dans les cellules de toutes les espèces qui ont besoin d'oxygène pour vivre – plantes, animaux ou bactéries. Grâce à l'analyse des molécules de cytochrome c chez plusieurs espèces, on a découvert que ces molécules, chez l'homme, ne différaient de celles des singes rhésus que par un seul acide aminé dans toute la chaîne. Il existait dix différences entre le cytochrome c de l'homme et celui du kangourou ; vingt et une différences entre l'homme et le thon ; et quelque quarante différences entre l'homme et une cellule de levure.

Avec l'aide de l'ordinateur, les calculs des biochimistes ont montré qu'il faut environ sept millions d'années pour qu'une modification dans un acide aminé se produise, et à partir de là, on peut estimer quand ont eu lieu des divergences entre divers types d'organismes. Il y a environ 2,5 milliards d'années, si l'on en croit les analyses effectuées sur le cytochrome c, que les organismes développés ont divergé des bactéries (c'est donc à peu près à cette époque qu'aurait existé une créature vivante pouvant être considérée comme l'ancêtre commun de tous les eucaryotes). De même, c'est il y a environ 1,5 milliard d'années que les végétaux et les animaux ont eu un ancêtre commun, et 1 milliard d'années que les insectes et les vertébrés ont eu un ancêtre commun. Il est donc clair que la théorie de l'évolution n'est pas seulement fondée sur les fossiles, mais également sur un large éventail de données géologiques, biologiques et biochimiques.

LA CINÉTIQUE DE L'ÉVOLUTION

Si les mutations dans la chaîne d'ADN qui entraînent des modifications au niveau de la séquence des acides aminés ne sont dues qu'au hasard, on pourrait admettre que l'évolution a suivi un cours plus ou moins

constant. Pourtant, il semblerait que l'évolution « accélère » à certaines périodes – quand de nouvelles espèces apparaissent soudain – comme le propose Gould avec sa notion d'évolution ponctuée, que j'ai mentionnée précédemment. Il est possible que le taux de mutations ait été plus élevé à certaines périodes de l'histoire de la Terre qu'à d'autres, et cette fréquence plus élevée de mutations explique peut-être l'apparition d'un grand nombre de nouvelles espèces ou l'extinction d'une multitude d'autres. Ou bien encore, quelques-unes de ces nouvelles espèces se sont avérées tellement mieux adaptées qu'elles ont concurrencé les plus anciennes jusqu'à les éliminer complètement.

L'un des facteurs de l'environnement qui favorisent le plus les mutations se trouve être les radiations d'énergie, et la Terre en est constamment bombardée de toutes parts. L'atmosphère en absorbe la quasi-totalité, mais même cette barrière est inefficace contre les radiations cosmiques. Se peut-il que les radiations cosmiques soient plus intenses à certaines périodes qu'à d'autres ?

On peut aborder cette question sous deux angles. Le champ magnétique terrestre fait diverger, dans une certaine mesure, les radiations cosmiques. Mais l'intensité de ce champ varie, et à certaines époques dont la fréquence est irrégulière, il peut même tomber à zéro. En 1966, Bruce Heezen a suggéré que ces périodes sont propices au bombardement massif de la surface du globe par les rayons cosmiques, qui entraîneraient un brusque accroissement du taux de mutations. Cette hypothèse donne certainement matière à réflexion, car il semblerait que la Terre soit sur le point d'entrer dans une telle période.

Certains astronomes ont également avancé l'hypothèse de l'apparition de supernovae à proximité de la Terre – c'est-à-dire assez proches du système solaire pour provoquer une élévation de l'intensité du bombardement cosmique à la surface de la Terre.

Les origines de l'homme

Pour James Ussher, un archevêque irlandais du XVIIe siècle, la création de l'homme (terme utilisé couramment pour englober les êtres humains des deux sexes, jusqu'à l'avènement du mouvement féministe dans les années soixante) remontait exactement à l'an 4004 av. J.-C.

Avant Darwin, bien peu de gens mettaient en doute l'interprétation biblique de l'histoire humaine. Le règne de Saül représente, dans les Écritures, la période historique la plus éloignée que l'on puisse déterminer avec un degré raisonnable d'exactitude ; on pense généralement que le premier roi d'Israël a vécu autour de 1025 av. J.-C. Ussher et d'autres érudits ont ainsi remonté la chronologie biblique et sont arrivés à la conclusion que l'homme et l'Univers qui l'entoure ne pouvaient être âgés de plus de quelques milliers d'années.

LES CIVILISATIONS ANTIQUES

L'histoire humaine, telle que les historiens grecs se la représentaient, n'était guère plus précise et ne remontait pas plus loin dans le passé que

le récit biblique. Elle n'a été véritablement documentée qu'à partir de 700 av. J.-C. environ. Au-delà n'existaient que de vagues légendes sur la guerre de Troie, vers 1200 av. J.-C., et d'autres encore plus imprécises évoquant une civilisation pré-hellénique sur l'île de Crète au temps d'un certain roi Minos. A part certains ouvrages écrits par des historiens dans des langues connues, avec tout ce que cela peut comporter de partialité et de distorsions, aucune source d'information sur la vie quotidienne dans les temps anciens n'a été disponible avant le XVIIIe siècle. C'est en 1738 qu'ont commencé les travaux d'excavation de Pompéi et d'Herculanum, deux cités enfouies sous les cendres du Vésuve lors d'une explosion volcanique en 79 apr. J.-C. Pour la première fois, les historiens ont compris quel profit ils pouvaient tirer des fouilles sur le terrain, marquant ainsi le début d'une nouvelle science, l'archéologie.

Au début du XIXe siècle, les archéologues ont eu un premier aperçu des civilisations antérieures aux récits des historiens grecs et hébreux. En 1799, au cours de la campagne d'Égypte du général Bonaparte, un de ses officiers, Boussard, a découvert une pierre gravée dans la ville de Rosette, située à l'une des embouchures du Nil. Sur la tablette de basalte noir figuraient trois inscriptions, une en grec, une en *hiéroglyphes* (« écriture sacrée »), c'est-à-dire la vieille forme d'écriture égyptienne, et la dernière sous une forme simplifiée de l'écriture égyptienne, appelée *démotique* (« du peuple »).

Le texte grec était un banal décret du temps de Ptolémée V, daté de l'équivalent du 27 mars 196 av. J.-C. De toute évidence, il devait s'agir de la traduction du même décret dans les deux autres langues (un peu comme les panneaux d'interdiction de fumer et autres avis officiels qui figurent souvent aujourd'hui en trois langues dans les lieux publics, particulièrement les aéroports). Les archéologues se frottaient les mains : ils allaient enfin pouvoir comprendre les manuscrits égyptiens, jusqu'alors indéchiffrables. Thomas Young a effectué un travail considérable qui a permis de « décoder » les hiéroglyphes ; c'est également à lui qu'on doit la théorie ondulatoire de la lumière (voir chapitre 8). Mais c'est Jean François Champollion, spécialiste français de l'Antiquité, qui est parvenu à résoudre l'énigme de *la pierre de Rosette*. Il est parti de l'idée que le copte, une langue pratiquée par certaines sectes chrétiennes d'Égypte, pourrait servir de guide pour déchiffrer les hiéroglyphes. En 1821, il a réussi à « déplomber » le code des hiéroglyphes et de l'écriture démotique, permettant ainsi la lecture de toutes les inscriptions découvertes dans les ruines de l'Égypte antique.

Une découverte presque identique a permis quelques années plus tard de déchiffrer les écrits mésopotamiens. Sur une haute falaise où se situaient les ruines du village de Béhistûn, dans l'Ouest de l'Iran, des chercheurs ont découvert un texte sculpté vers 520 av. J.-C., à la demande de l'empereur persan Darius. Il relatait la façon dont il avait pris le pouvoir après sa victoire sur un usurpateur ; pour s'assurer que tout le monde pourrait le lire, Darius l'avait fait graver en trois langues – le persan, le sumérien et le babylonien. Les textes sumériens et babyloniens étaient écrits sous forme de pictogrammes remontant à 3100 av. J.-C., gravés dans de l'argile au moyen d'un stylet ; ils avaient donné naissance à une écriture *cunéiforme,* utilisée jusqu'à la fin du Ier siècle apr. J.-C.

Un officier de l'armée britannique, Henry Creswicke Rawlinson, grimpa au sommet de la colline, nota tous les signes, et en 1846, après dix années de travail, en termina la traduction en s'aidant des dialectes locaux. Le

déchiffrement des manuscrits cunéiformes ouvrait la voie à l'étude historique des civilisations qui avaient façonné le berceau du monde entre le Tigre et l'Euphrate.

L'Égypte et la Mésopotamie devinrent les lieux privilégiés de nombreuses expéditions scientifiques ayant pour but de retrouver des tablettes et les restes des civilisations antiques qui y avaient fleuri. En 1854, un savant turc, Hurmuzd Rassam, découvrit les restes d'une bibliothèque de tablettes d'argile dans les ruines de Ninive, la capitale de l'Assyrie antique – une bibliothèque qui avait été constituée par son dernier roi, Assurbanipal, vers 650 av. J.-C. En 1873, l'Anglais George Smith, un assyriologue distingué, découvrit d'autres tablettes d'argile sur lesquelles figuraient les récits légendaires d'un déluge dont les détails correspondaient tellement à l'histoire de Noé qu'il apparut avec certitude que le premier livre de la Genèse était inspiré d'une légende babylonienne. Il est probable que les juifs eurent connaissance de ces légendes pendant leur captivité babylonienne à l'époque de Nabuchodonosor, un siècle après le règne d'Assurbanipal. En 1877, une expédition française envoyée en Iraq découvrit les restes de la civilisation qui avait précédé Babylone – celle des Sumériens dont j'ai parlé plus haut. Cette découverte faisait remonter l'histoire jusqu'aux périodes égyptiennes connues les plus reculées. En 1921, les restes d'une civilisation jusqu'alors totalement inconnue furent exhumés dans la vallée de l'Indus, une région appartenant aujourd'hui au Pakistan. Elle avait fleuri entre les années 2500 et 2000 av. J.-C.

Et pourtant, ni l'Égypte ni la Mésopotamie n'ont pu rivaliser en richesse avec les découvertes faites en Grèce sur les origines de la culture occidentale. En 1873, s'est peut-être déroulé l'épisode le plus passionnant de l'histoire de l'archéologie, lorsque le fils d'un ancien épicier allemand découvrit la plus célèbre de toutes les cités légendaires.

Depuis son plus jeune âge, Heinrich Schliemann s'était consacré à l'étude d'Homère. Alors que la plupart des historiens ne voyaient en l'*Iliade* que pure mythologie, Schliemann vivait et rêvait de la guerre de Troie. Il décida qu'il devait retrouver Troie, et aux prix d'efforts surhumains, se hissa de la condition de fils d'épicier à celle de millionnaire pour pouvoir financer son rêve. En 1868, à l'âge de quarante-six ans, il entama son périple. Il persuada le gouvernement turc de lui accorder la permission d'effectuer des fouilles en Asie Mineure ; puis, se basant seulement sur les maigres indices géographiques fournis par Homère, il se retrouva au sommet d'une petite colline près du village d'Hissarlik. Là, il persuada la population locale de l'aider à fouiller la colline de fond en comble. Les excavations, menées en dépit du bon sens, sans aucune rigueur scientifique, révélèrent une série de cités enfouies, chacune construite sur les ruines de la précédente. Et puis enfin, ce fut la réussite : il découvrit Troie – ou du moins la ville qu'il pensait être Troie. On sait maintenant que les ruines qu'il a appelées Troie sont en fait beaucoup plus anciennes que la cité d'Homère, mais Schliemann avait prouvé que les récits homériques n'étaient pas de simples légendes.

Grisé par le succès, Schliemann se rendit en Grèce et entreprit des fouilles à Mycènes, un village en ruines où Homère avait situé la puissante métropole d'Agamemnon, le grand stratège grec dans la guerre de Troie. Et de nouveau, Schliemann fit une découverte étonnante – les ruines d'une forteresse aux murs gigantesques, que nous savons maintenant avoir existé vers 1500 av. J.-C.

Ces succès répétés de Schliemann donnèrent l'idée à l'archéologue anglais Arthur John Evans d'entamer des fouilles sur l'île de Crète, où les légendes grecques situaient une puissante civilisation conduite par un certain roi Minos. Les efforts d'Evans se soldèrent, dans les années 1890, par la découverte des restes d'une civilisation brillante, caractérisée par des décorations somptueuses, qui remontait à plusieurs siècles avant la Grèce d'Homère. Là aussi, on y trouva des tablettes couvertes d'inscriptions, écrites dans deux langues, dont l'une, appelée linéaire B, fut déchiffrée dans les années cinquante. Elle était proche du grec, comme l'ont démontré les remarquables travaux de cryptographie et d'analyse linguistique du jeune architecte anglais Michael Vestris.

Au fur et à mesure qu'on découvrait de nouvelles civilisations antiques – celles des Hittites et des Mitanniens en Asie Mineure, de l'Indus aux Indes, etc. – il est apparu évident que l'histoire de la Grèce écrite par Hérodote et l'Ancien Testament des Hébreux témoignaient de civilisations à des stades relativement développés. Les cités les plus anciennes dataient d'au moins plusieurs milliers d'années, et l'existence préhistorique d'êtres humains aux modes de vie moins civilisés devait donc remonter à plusieurs milliers d'années auparavant.

L'ÂGE DE PIERRE

Les anthropologues trouvent commode de diviser l'histoire culturelle de l'homme en trois grandes périodes : l'âge de pierre, l'âge de bronze et l'âge de fer (c'est le poète latin Lucrèce qui a tout d'abord suggéré cette division, et le paléontologue danois Christian Jürgensen Thomsen l'a introduite dans la science moderne en 1834). Avant l'âge de pierre, il se peut qu'un « âge de l'os » ait existé, lorsque les hommes se servaient de cornes pointues, de dents aiguisées, et de fémurs en forme de massue, à une époque où les techniques du travail de la pierre n'avaient pas encore été mises au point.

Les âges de bronze et de fer sont bien sûr très récents ; dès que nous abordons les périodes précédant l'histoire écrite, nous sommes à l'âge de pierre. Ce que nous appelons *civilisation* (du mot latin signifiant « cité ») a peut-être commencé vers 8000 av. J.-C., quand les hommes se sont davantage consacrés à l'agriculture qu'à la chasse, ont appris à domestiquer les animaux, inventé la poterie et de nouveaux types d'outils, fondé des communautés permanentes, et adopté un style de vie sédentaire. On appelle cette période de transition âge de pierre, ou période néolithique, parce que des outils en pierre témoignant de techniques relativement sophistiquées la caractérisent. L'humanité existait donc à une époque bien antérieure à la date présumée de la Création telle qu'elle est décrite dans la Bible ; elle était en fait déjà « vieille » en ces temps-là.

Cette révolution du Néolithique semble avoir commencé au Proche-Orient, à la croisée des chemins de l'Europe, de l'Asie et de l'Afrique (où les âges de bronze et de fer allaient naître plus tard). De là, il semblerait qu'elle se soit étendue lentement au reste du monde en vagues successives. Il a fallu attendre 3000 av. J.-C. pour qu'elle atteigne l'Europe occidentale et l'Inde, 2000 av. J.-C. pour l'Europe du Nord et l'Extrême-Orient, et 1000 av. J.-C. ou plus tard pour l'Afrique centrale et le Japon. L'Afrique du Sud et l'Australie sont restées à l'âge de pierre jusqu'aux XVIIIe et XXe siècles.

La plus grande partie du continent américain en était encore au stade de la chasse lorsque les Européens sont arrivés au XVIe siècle, même si une civilisation très développée, vraisemblablement née des Mayas, fleurissait en Amérique centrale et au Pérou depuis les premiers siècles de l'ère chrétienne.

En Europe, on a découvert pour la première fois des témoignages de cultures prénéolithiques à la fin du XVIIIe siècle. En 1797, l'Anglais John Frere exhuma dans le Suffolk des outils de silex grossièrement taillés, et d'une facture trop primitive pour dater du Néolithique. Il les trouva enfouis sous un peu plus de quatre mètres de terre, indiquant ainsi une époque plus reculée si l'on tient compte d'une vitesse de sédimentation raisonnable. Dans la même strate, les outils côtoyaient des os d'animaux disparus. Des traces de plus en plus nombreuses du passé lointain de l'homme ont été découvertes, particulièrement par deux archéologues français du XIXe siècle, Jacques Boucher de Perthes et Édouard Armand Lartet. Ce dernier, par exemple, trouva une dent de mammouth sur laquelle une excellente reproduction de l'animal avait été méticuleusement gravée, sans aucun doute d'après un spécimen vivant. Le mammouth est une espèce d'éléphant à poil long qui a disparu bien avant le début du Néolithique.

Les archéologues se sont alors lancés activement à la recherche d'outils primitifs en pierre. Ils découvrirent qu'on pouvait les classer en un âge de pierre moyen (le Mésolithique) et un âge de pierre ancien (le Paléolithique). Le Paléolithique est divisé en périodes : basse, moyenne et haute. Les objets les plus anciens pouvant être considérés comme des outils (des *éolithes*) semblaient dater d'il y a environ un million d'années !

Quel genre de créature avait pu fabriquer les outils de l'âge de pierre ancien ? Il s'avéra que les hommes du Paléolithique étaient bien plus que des chasseurs. En 1879, un noble espagnol, le marquis de Sautuola, explora des cavernes qui avaient été découvertes quelques années auparavant – après avoir été bouchées par des glissements de roches successifs depuis la préhistoire – à Altamira dans le Nord de l'Espagne, près de la ville de Santander. Alors qu'il creusait le sol d'une de ces cavernes, sa fille âgée de cinq ans, qui était venue admirer son papa dans ses œuvres, se mit soudain à crier « Toros ! Toros ! ». Le père se releva, et vit sur les murs de la grotte les dessins très détaillés de plusieurs animaux aux couleurs vives.

Les anthropologues eurent beaucoup de mal à admettre que ces peintures très sophistiquées étaient effectivement l'œuvre d'un peuple primitif. Mais certains des animaux qui y figuraient appartenaient sans aucun doute possible à des espèces disparues. L'archéologue français Henri Edouard Prosper Breuil découvrit des exemples d'une facture similaire dans des grottes du Sud de la France. Cet ensemble de preuves a finalement convaincu les archéologues de l'exactitude des opinions de Breuil, et ils ont conclu que les artistes concernés avaient probablement vécu à la fin du Paléolithique, c'est-à-dire aux environs de 10 000 av. J.-C.

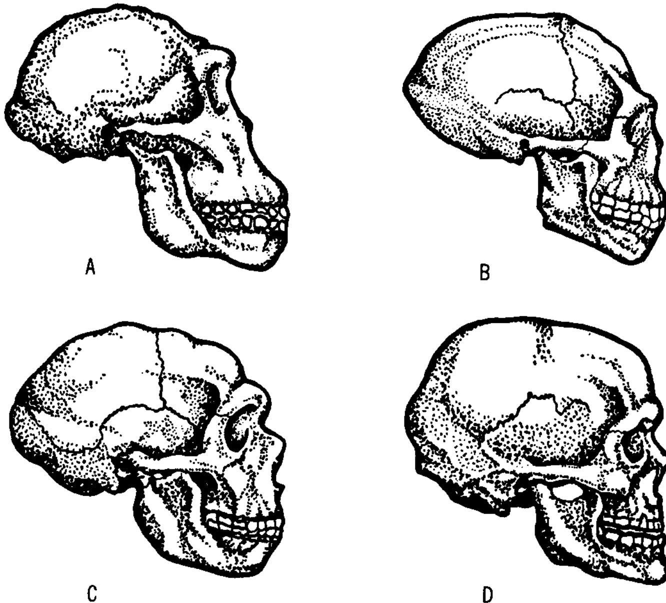
Certains détails concernant l'aspect physique de ces hommes du Paléolithique étaient déjà connus. En 1868, des ouvriers travaillant à la construction d'une voie de chemin de fer étaient tombés sur les squelettes de cinq êtres humains dans les cavernes de Cro-Magnon, dans le Sud-Ouest de la France. Il ne faisait aucun doute que ces squelettes appartenaient à des *Homo sapiens*, et pourtant certains d'entre eux, ainsi que d'autres

squelettes similaires découverts ailleurs peu après, semblaient avoir au moins 35 000 ou 40 000 ans, si l'on en croyait les études géologiques. On les appela *hommes de Cro-Magnon* (figure 16.4). Plus grand que l'homme moderne moyen, doté d'une grosse boîte crânienne, les artistes contemporains représentent l'homme de Cro-Magnon sous les traits d'un bel individu robuste, à l'aspect très moderne : il est fort probable qu'il n'aurait aucun mal à se reproduire avec les humains d'aujourd'hui.

Les êtres humains, tels que nous avons retracé leur histoire jusqu'à maintenant, ne vivaient pas sur toute la surface du globe comme aujourd'hui. Jusque vers 20 000 av. J.-C., ils se trouvaient confinés dans la « grande île du monde », c'est-à-dire l'Afrique, l'Asie et l'Europe. Ce n'est que plus tard que des hordes de chasseurs commencèrent à migrer par d'étroits passages marins sur le continent américain, en Indonésie, et en Australie. Il a fallu attendre 400 ans av. J.-C., et même plus tard, pour que de courageux Polynésiens traversent d'immenses étendues de l'océan Pacifique, sans boussole, et dans des embarcations guère différentes de canoës, pour coloniser les îles du Pacifique. Enfin, ce n'est pas avant le XXe siècle que le pied de l'homme a foulé le sol de l'Antarctique.

Mais afin de pouvoir retracer les destinées des peuplades préhistoriques à l'époque où elles étaient implantées sur une partie seulement du globe, il faut des moyens pour dater les événements, même grossièrement. Pour cela, les chercheurs ont mis au point plusieurs méthodes particulièrement ingénieuses.

Figure 16.4. Crânes reconstitués de (A) Zinjanthropus, (B) Pithecanthropus, (C) Néanderthal, (D) Cro-Magnon.



Les archéologues ont utilisé, par exemple, les anneaux des troncs d'arbres, une technique (la *dendrochronologie*) introduite en 1914 par l'astronome américain Andrew Ellicott Douglass. Ces anneaux sont très écartés pendant les étés pluvieux, lorsque le bois nouveau se forme, et peu écartés au cours des étés secs. Cette séquence est particulièrement frappante au cours des siècles. La séquence particulière des anneaux d'un bois ayant servi à la construction d'une habitation primitive peut être comparée à celle d'un arbre local « témoin », permettant ainsi de dater ce bois.

Un système identique peut être appliqué aux couches sédimentaires, ou *varves*, déposées été après été par la fonte des glaciers, comme par exemple en Scandinavie. Les étés chauds sont caractérisés par des couches épaisses, les étés frais par des couches minces ; de nouveau, chaque séquence est très nette. En Suède, ce système peut servir à dater des événements qui se sont produits jusqu'à il y a 18 000 ans.

Le chimiste américain Willard Frank Libby a développé une technique encore plus étonnante en 1946. Il est parti de la découverte en 1939 par le physicien américain Serge Korff du fait que le bombardement de l'atmosphère par les rayons cosmiques produit des neutrons. L'azote réagit avec ces neutrons en produisant du carbone 14 radioactif dans neuf réactions sur dix, et de l'hydrogène 3 radioactif dans la dixième réaction. Il résulte que l'atmosphère doit toujours contenir des traces faibles de carbone 14 (et des quantités encore plus faibles d'hydrogène 3). Libby a tenu le raisonnement que le carbone 14 radioactif créé dans l'atmosphère par les rayons cosmiques pénètre tous les tissus vivants par le biais du gaz carbonique, absorbé tout d'abord par les plantes puis assimilé par les animaux. Tant qu'une plante ou un animal vit, il ou elle continue d'ingérer des quantités de carbone radioactif, et le maintient à un niveau constant dans ses tissus. Mais quand l'organisme meurt et cesse d'absorber du carbone, la quantité de carbone radioactif présent dans les tissus commence à diminuer par le processus de la désintégration radioactive, à une vitesse déterminée par sa demi-vie de 5 600 années. Ainsi, n'importe quel morceau d'os, n'importe quel bout de charbon de bois ayant servi à faire un feu à une époque reculée, n'importe quel résidu organique, quel qu'il soit, peut être daté en mesurant sa teneur en carbone radioactif. Cette méthode est raisonnablement exacte pour des objets ayant au plus 30 000 ans, ce qui recouvre l'histoire archéologique depuis les civilisations antiques jusqu'à l'apparition de l'homme de Cro-Magnon. Pour couronner ce travail qui a donné naissance à l'*archéométrie*, Libby a reçu le prix Nobel de chimie en 1960.

Cro-Magnon n'est pas le premier homme primitif exhumé par les archéologues. En 1857, dans la vallée allemande du Néanderthal en Rhénanie, une fouille a mis en évidence un morceau de crâne, ainsi que des os longs qui paraissaient en gros d'origine humaine, mais le doute persistait néanmoins. Le crâne possédait un front à l'oblique très prononcée et des arcades sourcilières particulièrement proéminentes. Certains archéologues soutenaient qu'il s'agissait là des restes d'un être humain dont les os avaient été déformés par une maladie ; mais au fil du temps, on a retrouvé d'autres squelettes identiques, et on a pu se faire une idée plus précise et plus cohérente de ce qu'avait été l'homme de Néanderthal. Néanderthal était un petit bipède trapu, à l'allure générale voûtée ; les hommes ne mesuraient pas plus d'un mètre cinquante, les femmes un

peu moins. Le crâne pouvait loger un cerveau à peu près aussi gros que le nôtre (figure 16.4). Les dessins des anthropologues le représentent sous les traits d'une créature possédant une imposante cage thoracique, velu, aux sourcils épais et au menton pratiquement inexistant ; un véritable portrait de brute que l'on doit au paléontologue français Marcellin Boule, qui fut le premier à décrire un squelette quasi entier d'homme de Néanderthal en 1911. En réalité, il est probable qu'il était plus « humain » qu'on ne l'a décrit. Une étude récente du squelette décrit par Boule a démontré qu'il s'agissait d'un individu perclus d'arthrite. Un squelette normal révèle une créature beaucoup plus semblable à un homme actuel. En fait, rasé de prés, sortant de chez le coiffeur et de chez un bon tailleur, l'homme de Néanderthal pourrait sans doute descendre les Champs-Élysées sans détourner un seul regard.

Par la suite, des traces de l'homme de Néanderthal ont été retrouvées non seulement en Europe, mais aussi en Afrique du Nord, en Russie et en Sibérie, en Palestine, et en Iraq. Environ cent squelettes ont été découverts jusqu'à présent sur une quarantaine de sites, et il est possible que ces êtres humains vivaient encore il y a seulement 30 000 ans. Des restes de squelettes ressemblant à l'homme de Néanderthal ont été retrouvés dans des lieux encore plus éloignés les uns des autres ; il s'agit de l'homme de Rhodésie (aujourd'hui la Zambie) dans le Sud de l'Afrique en 1921, et de l'homme de Solo, exhumé aux abords de la rivière Solo à Java en 1931. On les considère comme des espèces distinctes du genre *Homo*, et on a donc baptisé respectivement chacun de ces trois types *Homo neanderthalensis, Homo rhodesiensis*, et *Homo solensis*. Mais certains anthropologues et spécialistes de l'évolution maintiennent que les trois devraient être classés dans la même espèce qu'*Homo sapiens*, en tant que *variétés* ou *sous-espèces* de l'homme. Des êtres humains que nous appelons *sapiens* vivaient à la même époque que l'homme de Néanderthal ; comme des formes intermédiaires ont également été retrouvées, ceci laisse à penser qu'ils ont pu se reproduire entre eux. S'il est possible de classifier l'homme de Néanderthal et ses cousins parmi les *sapiens*, notre espèce a près de 250 000 ans.

LES HOMINIDÉS

La parution de *L'origine des espèces* de Darwin a déclenché une gigantesque battue aux quatre coins du monde dans le but de retrouver nos ancêtres – ce que la presse populaire a appelé le *chaînon manquant* entre nous et nos prédécesseurs, théoriquement plus proches du singe. Cette entreprise, par la nature même des choses, ne pouvait être facile. Les primates sont des êtres doués d'une intelligence certaine, ce qui leur évite de tomber dans des situations propices à la fossilisation : on estime généralement qu'il y a une chance sur un quadrillion de retrouver un squelette de primate par hasard au cours d'une fouille.

Dans les années 1880, le paléontologue néerlandais Marie Eugène François Thomas Dubois se mit dans la tête qu'il fallait rechercher les ancêtres de l'homme dans les îles de la Sonde (aujourd'hui l'Indonésie), où les grands singes se propageaient toujours activement (et où il lui était facile de conduire des recherches car ces îles appartenaient alors aux Pays-Bas). Quelle ne fut pas sa surprise, alors qu'il travaillait à Java, la

plus peuplée des îles indonésiennes, de mettre la main sur une créature qui ressemblait à la fois à un singe et à un être humain ! Après une quête de près de trois ans, il avait trouvé le dessus d'un crâne plus gros que celui d'un singe mais plus petit que ceux reconnus généralement d'origine humaine. L'année suivante, il découvrait un fémur également de taille intermédiaire. Dubois baptisa son homme de « Java » *Pithecanthropus erectus* (« homme-singe droit ») (figure 16.4). Un demi-siècle plus tard, dans les années trente, un autre Néerlandais, Gustav von Koenigswald, exhuma d'autres os de *Pithecanthropus*, qui permirent d'établir l'allure générale d'une créature à petite tête et aux arcades sourcilières très proéminentes qui ressemblait vaguement à l'homme de Néanderthal.

Pendant ce temps, d'autres chercheurs avaient mis au jour, dans une grotte près de Pékin, des crânes, des mâchoires, et des dents d'un homme primitif qu'ils appelèrent bien sûr *homme de Pékin*. Après cette découverte, on s'aperçut qu'on avait déjà signalé de telles dents – dans une pharmacie de Pékin, où elles servaient à des fins médicinales. Le premier crâne intact fut découvert en décembre 1929, et le rapprochement établit des similitudes frappantes entre l'homme de Pékin et l'homme de Java. Celui-là vivait voici peut-être un demi-million d'années, connaissait le feu, et fabriquait des outils à partir d'os et de pierre. Finalement, on a retrouvé des fragments ayant appartenu à quarante-cinq individus, mais ils disparurent lors d'une évacuation précipitée des fossiles, provoquée par l'approche des troupes japonaises. En 1949, des archéologues chinois reprirent les fouilles, et les restes de quarante individus des deux sexes et d'âges différents sont maintenant répertoriés.

L'homme de Pékin s'appelle *Sinanthropus pekinensis* (« l'homme chinois de Pékin »), mais un examen plus approfondi de ces *hominidés* (créatures « ressemblant à l'homme »), au cerveau relativement peu développé, a démontré qu'il était peu fondé de classifier l'homme de Pékin et l'homme de Java dans des genres différents. Le biologiste germano-américain Ernst Walter Mayr pensait qu'il était incorrect de les placer dans un genre différent de celui des hommes d'aujourd'hui, si bien que l'on considère actuellement que l'homme de Pékin et l'homme de Java sont deux variétés de l'espèce *Homo erectus*, dont les premiers membres sont apparus sur la Terre il y a peut-être 700 000 ans.

Il est peu probable que l'espèce humaine soit née à Java, malgré l'existence là-bas d'un hominidé doté d'un petit cerveau. Pendant quelques années, on a supposé que l'humanité avait vu le jour au sein du vaste continent asiatique, là où avait vécu l'homme de Pékin ; mais au cours du XXe siècle, l'attention s'est tournée de plus en plus vers le continent africain qui, après tout, se trouve être le continent le plus riche en primates de tout genre, et particulièrement en primates développés. On y doit les premières découvertes intéressantes à deux chercheurs anglais, Raymond Dart et Robert Broom. Un jour du printemps 1924, des mineurs travaillant à la dynamite dans une carrière de calcaire près de la ville de Taungs en Afrique du Sud trouvèrent un petit crâne d'apparence très humaine. Ils l'envoyèrent à Dart, spécialiste en anatomie qui travaillait à Johannesburg. Dart l'identifia immédiatement comme provenant d'une créature mi-singe, mi-homme, et l'appela *Australopithecus africanus* (« singe austral d'Afrique »). A la publication de l'article annonçant sa découverte, les anthropologues pensèrent d'abord qu'il s'était trompé, et avait confondu un chimpanzé avec un homme-singe. Mais Broom, chasseur de fossiles

invétéré qui était depuis longtemps persuadé que les origines de l'homme se trouvaient en Afrique, prit le premier bateau pour Johannesburg et proclama haut et fort qu'*Australopithecus* représentait le spécimen le plus proche du chaînon manquant jamais découvert jusque-là.

Au cours des décades qui suivirent, Dart, Broom et bien d'autres anthropologues recherchèrent et découvrirent de nombreux autres échantillons d'os et de dents d'hommes-singes d'Afrique du Sud, ainsi que des massues utilisées par ces hominidés pour la chasse, des os d'animaux qu'ils avaient tués, et les grottes dans lesquelles ils avaient vécu. *Australopithecus* était une créature courtaude possédant un petit cerveau et un visage orné d'un mufle, sous bien des rapports moins humain que l'homme de Java. Mais les arcades sourcilières et les dents d'*Australopithecus* étaient d'apparence bien plus humaine que celles de *Pithecanthropus* ; en outre il marchait debout, utilisait des outils, et il est probable qu'il connaissait une forme primitive de langage. En résumé, *Australopithecus* formait une variété africaine d'hominidés ayant vécu il y a au moins un demi-million d'années et sans aucun doute plus primitifs qu'*Homo erectus*.

A première vue, il n'y avait aucune raison d'établir une priorité quelconque entre les hominidés d'Afrique et ceux d'Asie, mais la balance pencha finalement en faveur de l'Afrique grâce aux travaux de l'Anglais né au Kenya, Louis Seymour Bazett Leakey et de sa femme Mary. Avec patience et détermination, les Leakey ratissèrent les régions d'Afrique orientale propres à recéler des fossiles d'hominidés. La plus propice semblait être la gorge d'Olduvai, dans ce qui est actuellement la Tanzanie ; c'est là que, le 17 juillet 1959, les efforts de plus d'un quart de siècle de recherches furent couronnés par la découverte des fragments d'un crâne qui, une fois assemblé par Mary Leakey, s'avéra être la plus petite boîte crânienne d'hominidé jamais découverte. D'autres caractéristiques indiquaient cependant que cet hominidé était plus proche de l'homme que du singe, car il marchait debout, et les restes étaient entourés de petits outils faits de cailloux. Les Leakey baptisèrent leur trouvaille *Zinjanthropus* (« homme d'Afrique orientale », du mot arabe signifiant Afrique orientale) (figure. 16.4).

Il est peu probable que l'homme moderne descende en ligne directe de *Zinjanthropus*. Mais il semble que des fossiles encore plus âgés, datant d'il y a environ deux millions d'années, soient des postulants éventuels. Ceux-ci, à qui l'on a donné le nom d'*Homo habilis* (« homme agile »), mesuraient un peu plus d'un mètre trente, et leur pouce pouvait s'opposer aux autres doigts ; leur agilité (d'où leur nom) n'avait rien à envier à celle de l'homme actuel.

En 1977, l'archéologue américain Donald Johanson a découvert un fossile d'hominidé âgé de près de quatre millions d'années. Il a exhumé une quantité suffisante d'os pour assembler environ 40 % d'un individu entier. Il s'agit d'une petite créature d'environ un mètre de haut, caractérisée par des os très minces. Elle porte un nom scientifique, *Australopithecus afarensis*, mais on l'appelle vulgairement Lucie.

Le point le plus intéressant concernant Lucie, c'est qu'elle était totalement bipède – tout comme nous. Il semblerait que la principale caractéristique anatomique qui ait différencié les hominidés des singes soit effectivement cette faculté de marcher sur les deux jambes, à une époque où le cerveau des hominidés n'était guère plus gros que celui d'un gorille. Beaucoup pensent même, en fait, que le développement soudain et

remarquable de leur cerveau au cours de ce dernier million d'années est dû à cette faculté. Les membres antérieurs se sont retrouvés libres et sont devenus des mains délicates douées du sens tactile et capables de manipuler divers objets ; le flot d'informations atteignant le cerveau a renforcé les plus infimes développements de ce dernier, les rendant définitifs par le biais de la sélection naturelle.

Il est possible que Lucie soit en fait l'ancêtre de deux branches d'hominidés. Il y aurait d'un côté plusieurs australopithèques, dont le volume crânien variait entre 450 et 650 centimètres cubes, et qui ont disparu voilà près d'un million d'années ; de l'autre côté, nos ancêtres, les membres du genre *Homo*, comprenant *Homo abilis*, puis *Homo erectus* (dont la capacité crânienne variait entre 800 et 1 100 centimètres cubes), et finalement *Homo sapiens* (doté d'un cerveau allant de 1 200 à 1 600 centimètres cubes).

Naturellement, on trouve avant Lucie des fossiles d'animaux trop primitifs pour être qualifiés d'hominidés, et l'on s'approche de l'ancêtre commun des hominidés, c'est-à-dire du nôtre, et de celui des pongidés (ou singes), c'est-à-dire du chimpanzé, du gorille, de l'orang-outan, et de plusieurs espèces de gibbon.

Il y a *Ramapithecus*, dont le maxillaire supérieur a été retrouvé dans le Nord de l'Inde au début des années trente par G. Edward Lewis. Ce maxillaire était nettement plus « humain » que celui de n'importe quel autre primate, et date d'il y a environ trois millions d'années. En 1962, Leakey a découvert une espèce voisine âgée de quatorze millions d'années, comme l'ont démontré les études isotopiques.

En 1948, ce même chercheur a exhumé un fossile encore plus ancien (remontant à peut-être vingt-cinq millions d'années), baptisé Proconsul (« avant Consul ») en l'honneur de Consul, un chimpanzé du zoo de Londres. Il est probable que Proconsul est l'ancêtre commun des grands singes, le gorille, le chimpanzé, et l'orang-outan. Si l'on remonte encore plus loin dans le passé, un ancêtre commun à Proconsul et à *Ramapithecus* (et à un singe primitif qui a engendré le plus petit des grands singes actuels, le gibbon) doit très probablement exister. Une telle créature, la première de toutes les créatures ressemblant au singe, serait alors âgée d'environ quarante millions d'années.

L'HOMME DE PILTDOWN

Pendant de nombreuses années, les anthropologues se sont penchés avec une curiosité mêlée d'un certain embarras sur un fossile qui, à première vue, semblait être le chaînon manquant, mais tellement bizarre, tellement incroyable, qu'il fait encore parler de lui aujourd'hui. En 1911, non loin de Piltdown Common dans le Sussex, en Angleterre, des ouvriers qui construisaient une route trouvèrent les morceaux d'un vieux crâne au beau milieu de leur chantier. L'existence de ce crâne fut rapportée à un avocat du nom de Charles Dawson, qui l'amena chez un paléontologue, Arthur Smith Woodward, du British Museum. Le crâne possédait un front haut, des arcades sourcilières peu proéminentes ; il paraissait plus récent que l'homme de Néanderthal. Dawson et Woodward se mirent à la recherche d'autres fragments du squelette à l'endroit de la première découverte. Un beau jour, en présence de Woodward, Dawson trouva une mâchoire à peu

près au même endroit où l'on avait retrouvé les fragments de crâne. Elle avait le même aspect brun-rougeâtre que les autres fragments, et semblait donc appartenir à la même tête. Mais la mâchoire, contrairement à la partie supérieure, très humaine, du crâne, ressemblait à celle d'un singe ! Tout aussi étrange, les dents de la mâchoire, bien que semblables à celles d'un singe, étaient usées par la mastication comme celle d'un homme.

Woodward fut persuadé que ce spécimen mi-singe, mi-homme devait être une créature primitive possédant un cerveau comparativement bien développé et une mâchoire simiesque. Il présenta sa découverte au monde sous le nom d'*homme de Piltdown*, ou *Eoanthropus dawsoni* (« homme primitif de Dawson »).

L'homme de Piltdown commença à inspirer de sérieux doutes lorsque les anthropologues s'aperçurent que dans tous les cas où l'on avait retrouvé des fossiles de crânes avec leur mâchoire, le développement de cette dernière allait de pair avec celle du crâne. Finalement, au début des années cinquante, trois chercheurs anglais – Kenneth Oakley, Wilfrid Le Gros Clark et Joseph Sidney Weiner – décidèrent d'examiner l'hypothèse d'une supercherie. Et *c'était* effectivement une supercherie. La mâchoire, celle d'un singe contemporain, avait été placée là par des mains pour le moins malhonnêtes.

L'histoire de l'homme de Piltdown représente probablement l'exemple le plus connu et le plus embarrassant de scientifiques dupés pendant des années par un canular patent. Rétrospectivement, on peut s'étonner que des chercheurs soient tombés aussi facilement dans le panneau, mais la critique est un peu trop facile. Il faut se souvenir qu'en 1911, on connaissait peu de choses sur l'évolution des hominidés. Aujourd'hui, il est évident que pas un seul chercheur expérimenté ne se laisserait abuser par un tel canular.

Sur ce même sujet d'ossements de primate, une autre histoire étrange a trouvé un dénouement plus heureux. En 1935, von Koenigswald dénicha par hasard une énorme dent fossile, apparemment d'origine humaine, dans une pharmacie de Hong-Kong. Le pharmacien chinois pensait qu'il s'agissait là d'une « dent de dragon » aux précieuses vertus médicinales. Von Koenigswald fouina consciencieusement dans d'autres pharmacies et trouva quatre molaires semblables, avant que la Deuxième Guerre mondiale ne mette une fin temporaire à ses recherches.

L'origine humaine de ces dents laissait penser que des êtres gigantesques, d'une taille excédant peut-être deux mètres cinquante, avaient un jour existé sur Terre. Cette théorie recevait une certaine caution biblique, puisqu'on trouve dans les Écritures le verset suivant : « En ces jours-là il y avait des géants sur la Terre. » (Genèse, VI, 4.)

Entre 1956 et 1968, on retrouva néanmoins quatre mâchoires dans lesquelles on pouvait insérer ces dents. La créature, *Gigantopithecus*, est sans doute le plus grand primate ayant jamais existé, mais il ne fait non plus aucun doute qu'il s'agit d'un singe et non d'un hominidé, malgré l'apparence humaine de ses dents. Il ressemblait très probablement à un gorille, de deux mètres cinquante et plus debout, pesant près de trois cents kilos. Il était peut-être contemporain d'*Homo erectus* et partageait avec ce dernier les mêmes habitudes alimentaires (ce qui expliquerait la similitude des dents). Bien entendu, il a disparu voici au moins un million d'années et ne peut donc en aucun cas être à l'origine du verset biblique.

LES DIFFÉRENCES RACIALES

Il est important de souligner que l'aboutissement de l'évolution humaine est l'existence aujourd'hui d'une seule et unique espèce, autrement dit, bien qu'il ait existé un certain nombre d'espèces d'hominidés sur la Terre, une seule a survécu. Tous les hommes et toutes les femmes d'aujourd'hui, malgré des différences apparentes, sont des *Homo sapiens ;* et la différence qui existe entre les Noirs et les Blancs est à peu près du même ordre que celle qui existe entre des chevaux de différentes couleurs.

Et pourtant, depuis les origines de la civilisation, les hommes ont été plus ou moins intuitivement conscients de différences raciales et ont éprouvé, envers les autres, des sentiments pareils à ceux que l'on ressent en face d'étrangers : la curiosité, souvent le dédain, parfois même la haine. Mais le racisme a rarement produit les effets aussi tragiques et durables que le conflit entre les Blancs et les Noirs. (On appelle souvent les Blancs des Caucasiens, un terme tout d'abord utilisé par l'anthropologue allemand Johann Friedrich Blumenbach en 1775, qui avait commis l'erreur de croire que les représentants les plus purs de ce groupe se trouvaient dans le Caucase. Blumenbach a également classé les Noirs comme Éthiopiens et les Extrême-Orientaux comme Mongoliens, des termes que l'on entend encore parfois de nos jours.)

Le conflit racial entre Blancs et Noirs, autrement dit entre Caucasiens et Éthiopiens, est entré dans sa phase la plus aiguë au XVe siècle, lorsque des expéditions portugaises le long de la côte ouest de l'Afrique se lancèrent dans un trafic très lucratif d'esclaves. Au fur et à mesure qu'il prenait de l'importance et que les nations reposaient de plus en plus leur économie sur l'esclavage, on ressentit le besoin d'inventer de bonnes raisons pour justifier cette exploitation des Noirs : on évoqua les Écritures, la moralité sociale, et même la science.

Selon l'interprétation des négriers – une interprétation encore partagée de nos jours par beaucoup de gens – la Bible décrit les Noirs comme étant les descendants de Cham, et en tant que tels, comme les membres d'une tribu inférieure vouée à la malédiction de Noé : « Il sera pour ses frères l'esclave des esclaves. » (Genèse, IX, 25.) En fait, cette malédiction visait le fils de Cham, Canaan, et sa descendance, les Cananéens, qui furent réduits en esclavage par les juifs quand ces derniers firent la conquête de leur pays. Il ne fait aucun doute que le verset IX, 25 de la Genèse constitue un commentaire a posteriori, écrit par les Hébreux pour justifier l'exploitation des Cananéens. De toute façon, ce qu'il est important de retenir, c'est que la référence ne s'adresse qu'aux seuls Cananéens, et que ces derniers étaient certainement blancs. Ce sont les esclavagistes qui ont déformé les textes bibliques, avec toutes les conséquences dramatiques que cela a pu avoir au cours des siècles, pour défendre leur conduite vis-à-vis des Noirs.

Les racistes « scientifiques » contemporains se sont basés sur des arguments encore plus spécieux. Pour eux, les Noirs sont des êtres inférieurs car, de toute évidence, ils ont atteint un stade bien moins évolué que celui des Blancs. N'est-il pas utile de rappeler, par exemple, qu'une peau noire et des narines écartées évoquent immédiatement le singe ? Hélas, pour les racistes pseudo-scientifiques, ce genre d'arguments se retourne bien vite contre eux. Il s'avère que les Noirs forment le groupe humain le moins poilu ; de ce point de vue, et du fait que leurs cheveux

sont crépus et laineux, au lieu d'être longs et droits, ils sont plus éloignés du singe que ne l'est l'homme blanc ! On peut dire la même chose de l'épaisseur des lèvres des Noirs ; elles ressemblent moins à celles du singe que les lèvres minces des Blancs.

Qu'il suffise de dire une bonne fois pour toutes que tout effort déployé pour établir une quelconque classification des différents groupes d'*Homo sapiens* le long de l'échelle de l'évolution ne peut être que voué à l'échec. L'humanité n'est composée que d'une seule et même espèce, et jusqu'à présent les variations dues à la sélection naturelle n'ont été que superficielles.

La peau sombre des habitants des régions tropicales et subtropicales de la Terre protège évidemment des coups de soleil. La peau claire des Européens du Nord aide à l'absorption d'autant de rayons ultraviolets que possible provenant d'un soleil timide, pour fabriquer une quantité suffisante de vitamine D à partir des stérols de la peau. Les yeux bridés des Eskimos et des Mongols sont précieux pour survivre dans un environnement où la réverbération du soleil sur la neige ou sur les sables du désert est intense. La forme du nez des Européens avec ses narines étroites sert à réchauffer l'air froid des hivers septentrionaux. Et la liste est inépuisable.

Avec sa tendance à faire de sa planète un seul monde, aucune différence essentielle dans la constitution d'*Homo sapiens* ne s'est produite dans le passé, et il est encore moins probable qu'il s'en produise dans le futur. Le mélange des races nivèle l'héritage humain. L'homme noir américain en offre le plus bel exemple. Malgré les barrières sociales contre les mariages mixtes, on estime que du sang blanc coule dans près des trois quarts de la population noire aux États-Unis. Il est probable qu'à la fin de ce siècle, il n'y aura plus de Noirs de race « pure » en Amérique du Nord.

GROUPES SANGUINS ET RACES

Il n'en est pas moins vrai que les anthropologues s'intéressent vivement aux races, au premier chef comme guide pour expliquer les migrations des premiers êtres humains. Il n'est pas facile d'identifier des races précises. La couleur de la peau, par exemple, est un marqueur qui n'offre que peu d'intérêt ; l'aborigène d'Australie et le Noir d'Afrique ont tous les deux une peau sombre mais ne sont pas plus apparentés l'un à l'autre que chacun ne l'est à un Européen. La forme de la tête – *dolichocéphale* (long) contre *brachycéphale* (large), des termes introduits en 1840 par l'anatomiste suédois Anders Adolf Retzius – n'est guère plus utile, bien qu'une classification des Européens en sous-groupes ait été effectuée en suivant ce schéma. Le rapport entre la longueur et la largeur de la tête multiplié par 100 (l'*index céphalique*) a été utilisé pour classer les Européens en nordiques, alpins, et méditerranéens. Néanmoins, les différences d'un groupe à l'autre sont minimes, et la fourchette à l'intérieur de chacun d'eux est très large. De plus, certains facteurs externes, tels que les carences en vitamines, le type de berceau dans lequel un enfant dort, etc., ont une influence non négligeable sur la forme du crâne.

Mais les anthropologues ont découvert un excellent marqueur de la race : les groupes sanguins. Le travail du biochimiste William Clouser Boyd, de

l'université de Boston, est remarquable en la matière. Il a observé que nous héritons notre groupe sanguin par un processus simple et bien établi, qu'il n'est pas altéré par l'environnement, et qu'il apparaît à des fréquences bien distinctes d'une race à l'autre.

L'Indien d'Amérique du Nord en est un bon exemple. Certaines tribus sont presque toutes du groupe O; d'autres sont O mais avec une grande proportion de A ; et on ne trouve pratiquement pas d'Indiens à sang B ou AB. Un Indien américain, qui se révèle être B ou AB, est presque certain d'avoir des Européens dans ses ascendants. Quant aux aborigènes australiens, ils sont presque tous O et A, le groupe B étant pratiquement inexistant. Mais ils se distinguent de l'Indien américain par une fréquence plus élevée du groupe M, découvert récemment, et plus faible du groupe N, alors que ces fréquences sont inversées chez l'Indien américain.

En Europe et en Asie, où la population est plus mélangée, les différences entre peuples sont plus faibles, mais encore perceptibles. Par exemple, à Londres, 70 % de la population a du sang O ; 26 % du sang A ; et 5 % du sang B. Mais à Kharkov, en Union Soviétique, la distribution correspondante est la suivante : 60, 25, 15. En règle générale, le pourcentage de B augmente plus on voyage vers l'est de l'Europe, et l'on atteint un maximum de 40 % en Asie centrale.

Il est évident que les gènes qui déterminent le groupe sanguin révèlent les traces encore fraîches des anciennes migrations. L'infiltration du gène B en Europe est peut-être révélatrice, d'une façon atténuée, de l'invasion des Huns au v[e] siècle et des Mongols au XIII[e]. Des études hématologiques de ce type, menées en Extrême-Orient, indiquent la possibilité d'une infiltration assez récente du gène A au Japon par le sud-ouest et du gène B en Australie par le nord.

L'Espagne offre également une preuve intéressante et inattendue de très vieilles migrations en Europe, comme a pu le démontrer une étude menée sur la distribution des groupes Rh. Ces derniers portent leur nom d'après la réaction hématologique aux antisérums produits contre les globules rouges du sang d'un singe rhésus. Il existe au moins huit allèles du gène responsable ; sept d'entre eux sont *Rh positif*, et le huitième, récessif, est *Rh négatif*, car il n'est dominant que lorsqu'un individu reçoit cet allèle de ses deux parents. Aux États-Unis, environ 85 % de la population est Rh positif, ce qui laisse 15 % de Rh négatif. On retrouve cette même proportion dans la plupart des pays européens. Mais, curieusement, les Basques du Nord de l'Espagne forment un cas à part avec environ 60 % de Rh négatif et 40 % de positif. Et il est intéressant de noter que les Basques ont une langue qui n'a aucun rapport avec les autres langues européennes.

On peut donc en tirer la conclusion que la population basque représente les vestiges d'une invasion préhistorique de l'Europe par un peuple Rh négatif. Une vague d'invasions par des tribus Rh positif les a probablement enclavés ultérieurement dans leur refuge montagneux à l'ouest de l'Europe, où ils demeurent le dernier groupe important de survivants des *Européens primitifs*. Le faible résidu de gènes Rh négatif, que l'on retrouve dans le reste du continent et chez les descendants américains des colons européens, apporte un témoignage de ces événements lointains.

Les peuples d'Asie, les Noirs d'Afrique, les Indiens des États-Unis et les aborigènes d'Australie sont presque tous Rh positif.

L'avenir de l'humanité

Essayer de prédire l'avenir du genre humain comporte bien des risques, et il est préférable de laisser ce soin aux mystiques et aux écrivains de science-fiction (bien qu'après tout, je fasse partie de ces derniers). Mais nous pouvons être presque certains d'une chose : à moins d'une quelconque catastrophe mondiale – comme une guerre atomique généralisée, ou une attaque massive d'extra-terrestres, ou bien encore une pandémie dévastatrice de maladies d'un genre nouveau – la population de la Terre va augmenter rapidement. Elle est aujourd'hui cinq fois ce qu'elle était il y a seulement deux siècles. Des estimations portent le nombre total d'êtres humains ayant vécu pendant une période de 600 000 ans à 77 milliards. Si ces chiffres sont exacts, cela signifie que presque 6 % de tous les hommes qui ont jamais existé vivent à l'heure actuelle. Et la population mondiale ne cesse de s'accroître à une vitesse proprement phénoménale.

Comme nous ne possédons aucun recensement sur les populations anciennes, nous ne pouvons qu'établir des estimations grossières sur la base de ce que nous savons des conditions de vie qui régnaient en ces temps-là. Certains écologistes ont estimé que le niveau des ressources alimentaires dans la période précédant l'avènement de l'agriculture – c'est-à-dire provenant de la chasse, de la pêche, de la cueillette des fruits sauvages et des noix, etc. – n'aurait pu subvenir aux besoins d'une population excédant vingt millions d'individus, et que, très vraisemblablement, la population vivant à l'époque du Paléolithique ne dépassait pas au plus le tiers ou la moitié de ce chiffre. Donc, aussi récemment que 6 000 ans av. J.-C., elle ne devait pas représenter plus de six à dix millions d'âmes – moins que les habitants d'une seule métropole actuelle comme Shanghai ou Mexico. Lorsque Christophe Colomb a découvert l'Amérique, les Indiens qui vivaient sur le territoire de ce qui est maintenant les États-Unis ne devaient pas dépasser 250 000 – comme si la population de Dayton, dans l'Ohio, était répartie sur toute la surface du continent.

L'EXPLOSION DE LA POPULATION

Le premier grand bond de la population mondiale a eu lieu lors de la révolution du Néolithique, avec les débuts de l'agriculture. Le biologiste anglais Julian Sorrell Huxley (le petit-fils du Huxley qui avait défendu Darwin) estime qu'à cette époque le taux de croissance a commencé à augmenter à un rythme qui faisait doubler la population tous les 1 700 ans environ. A l'aube de l'âge de bronze, la Terre devait compter environ 25 millions d'individus ; au début de l'âge de fer, 70 millions ; au seuil de l'ère chrétienne, 170 millions, dont un tiers entassés à l'intérieur de l'Empire romain et un autre tiers dans l'Empire chinois, et le reste éparpillé aux quatre coins du globe. En 1600, la population de la Terre atteignait peut-être 500 millions d'habitants, un chiffre bien moins élevé que la population actuelle de la seule Inde.

C'est alors que le taux de croissance a commencé à s'emballer, et la population a littéralement explosé. Les explorateurs ouvraient à la colonisation des Européens plusieurs dizaines de millions de kilomètres carrés de terres vierges sur de nouveaux continents. La révolution

industrielle, au XVIIIe siècle, accéléra la production de ressources alimentaires et la reproduction des hommes. Même des pays arriérés comme la Chine ou l'Inde furent affectés par cette explosion. Dès lors, la population mondiale ne doublait plus en l'espace de presque deux millénaires, mais en moins de deux siècles. Elle passa de 500 millions en 1600 à 900 millions en 1800. Et depuis, elle a augmenté encore plus vite. En 1900, elle atteignait 1,6 milliard d'habitants. Dans les sept premières décades du XXe siècle, elle a grimpé à 3,6 milliards, malgré deux guerres mondiales.

En 1970, la population mondiale augmentait à un rythme de 220 000 individus par jour, c'est-à-dire 70 millions par an, équivalant à un accroissement annuel de l'ordre de 2 % (alors qu'il n'était que de 0,3 % en 1650). A ce rythme, la population de la Terre doublera en à peu près trente-cinq ans, et dans certaines régions, comme l'Amérique latine, ce doublement sera encore plus rapide.

A l'heure actuelle, les chercheurs qui étudient ce phénomène réexaminent avec une ferveur renouvelée les thèses de Malthus, qui ont mauvaise presse depuis qu'il les a formulées en 1798. Comme je l'ai dit précédemment, Thomas Robert Malthus soutenait, dans son *Essai sur les populations,* que ces dernières tendent toujours à s'accroître plus vite que les ressources alimentaires, entraînant invariablement des famines périodiques et des guerres. Malgré ses prédictions, la population mondiale a augmenté à grands pas, et sans rencontrer de problème majeur au cours de ce dernier siècle et demi. Mais ne serait-ce en fait que reculer pour mieux sauter ? Nous pouvons nous féliciter de la présence, à la surface du globe, de vastes étendues de terres arables. Mais ces terres se font aujourd'hui de plus en plus rares. La grande majorité de la population mondiale est sous-alimentée, et il nous faudra encore faire des efforts considérables pour remédier à ce phénomène chronique. Il est certain que les mers pourraient être exploitées de manière plus rationnelle : leur rendement en ressources alimentaires pourrait être ainsi grandement amélioré. On peut encore étendre l'utilisation des engrais chimiques. Les pesticides, si l'on en fait un usage raisonnable, réduiront les pertes encore causées par les insectes dans de nombreuses régions. Et il existe également beaucoup de moyens pour venir directement en aide aux cultures. Des hormones végétales telles que la *gibberelline* (étudiée par les biochimistes japonais avant la Deuxième Guerre mondiale, et introduite dans les pays occidentaux dans les années cinquante) peuvent stimuler la croissance des plantes, tandis que des antibiotiques, en petites quantités, accélèrent la croissance animale (peut-être en détruisant la flore intestinale qui, autrement, concurrence l'organisme en détournant à son profit une bonne partie des aliments qui transitent par les intestins, et en supprimant de petites infections qui affaiblissent cet organisme). Néanmoins, avec toutes ces nouvelles bouches qu'il faut nourrir et qui se multiplient à une vitesse impressionnante, il faudra déployer des efforts considérables pour simplement maintenir la population dans son état actuel, aussi déplorable soit-il, puisque quelque trois cents millions d'enfants de moins de cinq ans, à travers le monde, sont sous-alimentés au point de souffrir de lésions cérébrales incurables.

Même une ressource aussi banale (et, jusqu'à récemment, négligée) que l'eau commence à devenir rare. Le monde consomme actuellement près de huit milliards de mètres cubes d'eau par jour ; et bien que la pluie (c'est

elle qui fournit en ce moment la majeure partie de cette eau douce) qui tombe sur le globe quotidiennement représente cinquante fois ce chiffre, ce n'est qu'une faible proportion de cette eau que l'on peut facilement récupérer. Aux États-Unis, où l'on consomme 1,4 milliard de mètres cubes, la plus grande quantité par tête d'habitant de tous les pays du monde, ce sont presque 10 % des chutes de pluie totales qui sont consommées d'une manière ou d'une autre.

Cette situation explique pourquoi les lacs et les rivières deviennent de plus en plus fréquemment des pommes de discorde. La dispute qui oppose la Syrie et Israël au sujet du Jourdain, et celle qui met aux prises l'Arizona et la Californie à propos du fleuve Colorado, fournissent deux excellents exemples. On creuse des puits de plus en plus profonds ; et dans bien des parties du monde, le niveau des eaux sous la surface du sol baisse dangereusement. Pour tenter de préserver l'eau douce, on a utilisé l'alcool hexadécyclique pour recouvrir la surface des lacs et des réservoirs dans certaines régions d'Australie, d'Israël, et d'Afrique orientale. Ce produit se répand à la surface de l'eau sous forme d'une pellicule incroyablement fine, puisqu'elle n'est constituée que d'une seule molécule, réduisant ainsi l'évaporation sans polluer l'eau. (Bien sûr, l'augmentation de la pollution par les égouts et les déchets industriels est un autre facteur qui influe directement sur les quantités d'eau douce.)

Il ne fait plus guère de doute qu'un jour prochain, il sera nécessaire d'obtenir de l'eau douce à partir des océans ; ces derniers en offrent, du moins dans un avenir immédiat, des quantités quasi illimitées. Deux méthodes ont jusqu'à présent obtenu les meilleurs résultats pour dessaler l'eau de mer : la distillation et la congélation. De plus, on fait actuellement des expériences sur des membranes qui laisseront les molécules d'eau passer, mais barreront le chemin aux différents ions. Ce problème est d'une telle importance que les États-Unis et l'Union Soviétique discutent d'une stratégie commune, à une époque où la coopération entre les deux Grands est, dans bien d'autres domaines, extrêmement difficile à instaurer.

Mais soyons aussi optimistes que possible et imaginons que l'ingéniosité humaine soit sans limites. Supposons que, les miracles de la technologie aidant, nous augmentions le rendement de nos activités sur Terre par un facteur dix ; supposons que nous arrivions à extraire les métaux de l'Océan, que nous transformions tout le Sahara en un gigantesque champ de pétrole, que nous trouvions du charbon en Antarctique, que nous maîtrisions l'énergie solaire, et que nous développions la fusion nucléaire. Si nous parvenons à faire tout cela, et si le taux d'expansion de la population humaine continue sur sa lancée actuelle, toute notre science et tous nos progrès technologiques ne nous empêcheront pas, tel Sisyphe, de mener un combat sans espoir.

Si vous n'êtes pas encore tout à fait convaincus du bien-fondé de ces arguments, certes un peu pessimistes, considérons un instant les puissances de la progression géométrique. On pense généralement que la quantité totale de biomasse sur la Terre est actuellement égale à 2×10^{19}. S'il en est ainsi, la masse totale de l'humanité en 1970 était d'environ 1/100 000 de la biomasse sur Terre.

Si la population globale continue à doubler tous les trente-cinq ans (comme c'était le cas à cette époque), elle aura augmenté par un facteur 100 000 en l'an 2570. Il est probable qu'il sera extrêmement difficile d'accroître ainsi la masse vivante que la Terre peut supporter (bien qu'une

espèce puisse toujours se multiplier au détriment des autres). Dans ce cas, en 2570, la masse de l'humanité formera à elle seule toute la vie sur Terre, et nous en serons réduits à la pratique du cannibalisme pour que certains puissent survivre.

Même si nous pouvions produire des aliments artificiels à partir de la matière inorganique, au moyen des levures par exemple, ou de l'*hydroponique* (la culture de plantes dans des solutions chimiques), etc., aucun progrès concevable ne saurait pallier la croissance inexorable d'une population qui double tous les trente-cinq ans. A ce rythme, elle atteindrait 630 000 milliards d'individus ! Il ne resterait alors sur Terre que des places debout, puisque chaque personne disposerait d'un peu moins d'un mètre carré pour s'ébattre, qu'elle vive au Groënland ou dans l'Antarctique. En fait, si l'on pouvait imaginer que l'espèce humaine continue de se multiplier ainsi à ce même rythme, en 3550 la masse totale des tissus humains serait égale à la masse de la Terre.

Si certains voient dans l'émigration dans l'espace une possibilité de se tirer de ce mauvais pas, qu'ils réfléchissent un peu au fait que, en imaginant qu'il existe mille milliards d'autres planètes habitables dans l'Univers, et qu'on puisse effectivement y transporter tous les gens qui le désireraient, au rythme actuel d'accroissement toutes ces planètes seraient surpeuplées dès l'an 5000. En 7000, la masse de l'humanité serait égale à la masse de l'Univers connu tout entier !

De toute évidence, la race humaine ne peut pas continuer à augmenter à son rythme actuel pendant très longtemps, quoi qu'il advienne des questions de ressources alimentaires, d'eau, de minéraux et d'énergie. Je ne dis pas qu'elle ne « continuera pas » ou qu'elle ne « devrait pas continuer » ; je dis tout simplement qu'elle ne « peut pas ».

De fait, ce ne sont pas seulement les chiffres qui limiteront notre croissance, si elle continue à cette vitesse. Et ce n'est pas seulement le fait qu'à chaque minute qui passe, il y a plus d'hommes, plus de femmes et plus d'enfants ; mais c'est surtout que chaque individu (en moyenne) utilise une partie des ressources non renouvelables de cette Terre, dépense plus d'énergie, et produit à chaque instant plus de déchets et de pollution. Alors que la population doublait tous les trente-cinq ans, l'utilisation de l'énergie, en 1970, augmentait à un tel rythme qu'en trente-cinq ans, elle-même ne doublerait pas, mais serait multipliée par un facteur sept.

Ce besoin aveugle de gâcher et d'empoisonner de plus en plus chaque année notre environnement nous conduit inexorablement vers notre perte. Par exemple, la fumée qui provient de la combustion du charbon et du pétrole est relâchée librement dans l'atmosphère par les cheminées de nos logements et de nos usines, comme le sont les déchets chimiques gazeux de nos immenses complexes industriels. Les automobiles, par centaines de millions, libèrent des vapeurs d'essence et des produits provenant de sa dégradation et de son oxydation, sans même parler d'oxyde de carbone ou de plomb. Les oxydes de soufre et d'azote (produits soit directement, soit par oxydation ultérieure causée par les rayons ultraviolets du soleil), ainsi que d'autres substances, font rouiller les métaux, fragilisent les matériaux de construction, rendent le caoutchouc cassant, endommagent les récoltes, causent et aggravent les désordres respiratoires, constituent même l'une des causes majeures du cancer du poumon.

Lorsque les conditions atmosphériques sont telles que l'air stagne au-dessus d'une ville pendant un certain temps, les polluants s'accumulent,

contaminent sérieusement cet air et provoquent la formation d'un brouillard qui ressemble à de la fumée (le *smog*), décrit pour la première fois à Los Angeles, mais qui existait bien avant dans de nombreuses autres villes et qui gagne maintenant presque toutes les grandes métropoles. Dans les cas les plus dramatiques, il peut causer la mort de milliers de personnes qui, de par leur âge ou à cause des maladies dont elles souffrent, ne peuvent tolérer cette agression supplémentaire sur leurs poumons. De tels désastres se sont produits à Donora, Pennsylvanie, en 1948, et à Londres en 1952.

Les sources d'eau douce sur la Terre sont polluées par les déchets chimiques, et il arrive que l'une de ces substances fasse la une des journaux. Le cas s'est produit en 1970, quand on a découvert que des déchets contenant du mercure étaient jetés sans aucune précaution dans les mers du globe et qu'on les retrouvait dans l'organisme des produits de la pêche, parfois en quantité alarmante. A ce rythme, au lieu de trouver dans les océans une source inépuisable de nourriture, il est fort possible en fait que nous soyons en train de les empoisonner.

L'utilisation aveugle de pesticides quasi indestructibles conduit à leur absorption d'abord par les plantes, puis par les animaux. A cause de l'empoisonnement qui en résulte, certains oiseaux ont de plus en plus de mal à frabriquer des coquilles normales pour protéger leurs œufs, tant et si bien qu'en attaquant les insectes, nous en arrivons à mettre en danger l'existence du faucon pèlerin.

La plupart des prétendus progrès technologiques, sur lesquels nous nous précipitons tête baissée pour doubler un concurrent et faire plus de profits, peuvent être la cause des pires problèmes. Depuis la Deuxième Guerre mondiale, les détergents synthétiques ont remplacé le savon. Divers phosphates entrent dans leur fabrication, qui lorsqu'ils se retrouvent dans l'eau, favorisent la croissance de certains micro-organismes ; ceux-ci, hélas, épuisent les ressources en oxygène – entraînant la mort d'autres organismes vivant dans les mers. Ces changements nuisibles dans les habitats aquatiques (l'*eutrophication*) amènent un vieillissement prématuré des Grands Lacs, par exemple – en particulier le lac Érié, qui est peu profond – et raccourcissent leur durée de vie de plusieurs millions d'années. C'est ainsi que le lac Érié va peut-être devenir le marais Érié, et que les marécages des Everglades vont complètement s'assécher.

Les espèces vivantes sont interdépendantes. Il existe des cas typiques comme l'interrelation entre les plantes et les abeilles, dans laquelle les plantes sont pollinisées par les abeilles, et les abeilles nourries par les plantes ; et il y a mille autres cas moins évidents. Chaque fois que la vie est plus facile ou plus difficile pour une certaine espèce, ce sont des dizaines d'autres qui sont affectées – parfois par des voies bien mystérieuses. L'étude de cette interconnexion des formes de vie, l'*écologie*, fait maintenant parler d'elle, car nombreux sont les cas où les hommes, dans un souci égoïste de profit à court terme, ont tellement altéré la structure écologique qu'ils ont provoqué des catastrophes. Il est clair que nous devons agir en prenant bien plus de précautions que nous ne l'avons fait jusqu'à présent.

Même une activité apparemment aussi innocente que le lancement de fusées doit nous faire réfléchir. A elle seule, une grosse fusée peut injecter quelque cent tonnes de gaz d'émission dans l'atmosphère à des altitudes excédant cent kilomètres. De telles quantités de gaz modifient très probablement les propriétés de la haute atmosphère, augmentant ainsi le risque de perturber le climat de manière imprévisible. Dans les années

soixante-dix, on a vu apparaître des avions de ligne supersoniques qui traversent la stratosphère à près de deux fois la vitesse du son. Leurs détracteurs dénoncent non seulement le « bang » qu'ils provoquent quand ils franchissent le mur du son, mais aussi les changements de climat que la pollution causée par leurs réacteurs risque d'entraîner.

Un autre facteur qui rend encore plus dramatique tout accroissement de population se trouve être la distribution inégale des êtres humains à la surface du globe. Partout, la tendance est à une accumulation de la population dans les régions urbaines. Aux États-Unis, bien que la population augmente de plus en plus, non seulement certains États agricoles ne sont pas affectés par cette explosion, mais en fait leur population diminue. On estime que la population urbaine sur la Terre ne double pas tous les trente-cinq ans, mais tous les onze ans. A ce rythme en 2005, quand la population du monde aura doublé, la population des villes aura été multipliée par un facteur 9.

Ce problème est trés sérieux. Nous assistons déjà à la désagrégation des structures sociales – une désagrégation qui est plus particulièrement notable justement dans les pays développés où l'urbanisation est la plus avancée. Elle est plus marquée dans les grandes villes, et spécialement dans les quartiers les plus denses. Il ne fait aucun doute que si l'on entasse les êtres humains au-delà d'un certain seuil critique, de nombreuses formes de comportement pathologique apparaissent brusquement. On a pu le démontrer au cours d'expériences chez le rat, et il suffit de lire les journaux ou de faire appel à sa propre expérience pour être convaincu que le même phénomène se produit chez l'être humain.

Il paraît donc évident que, *si rien n'est fait pour changer le cours actuel des choses,* les structures sociales et technologiques seront dans un état avancé de délabrement d'ici au milieu du siècle prochain, ce qui entraînera des conséquences incalculables. Pris de folie, les êtres humains auront peut-être alors envie de déclencher la catastrophe ultime, la guerre nucléaire. Mais *peut*-on changer le cours actuel des choses ?

Il est clair que pour changer ce cours, il faudra faire d'immenses efforts et abandonner bien des idées préconçues. Depuis les débuts de l'histoire, l'homme a vécu dans un monde où la vie était courte et où beaucoup d'enfants mouraient à peine nés. Pour que la tribu puisse se perpétuer, les femmes devaient porter autant d'enfants qu'elles le pouvaient. C'est pour cette raison que l'on vouait un culte à la maternité, et que toute attaque menée contre la natalité était impitoyablement repoussée. Le statut des femmes était réduit à celui de machines à fabriquer et élever les enfants. Les mœurs sexuelles étaient sévèrement codifiées et ne permettaient que les actes visant à la reproduction ; tout le reste n'était que perversion et péché.

Mais nous vivons aujourd'hui dans un monde surpeuplé. Si nous voulons échapper à une catastrophe, la maternité doit devenir un privilège accordé avec parcimonie. Nous devons réexaminer nos idées sur le sexe et sa connexion avec la reproduction.

Là encore, les problèmes du monde – les problèmes vraiment sérieux – nous affectent tous, où que nous vivions. Les dangers que présentent la surpopulation, la pollution, l'appauvrissement des ressources naturelles, le risque de guerre nucléaire, affectent tous les pays, et on ne trouvera de solutions valables que lorsque toutes les nations du monde accepteront de coopérer. Cela veut dire qu'aucune nation ne peut faire cavalier seul,

sans penser aux autres ; les nations ne peuvent plus prendre leurs décisions en fonction du mythe de la « sécurité nationale », selon lequel un événement bénéfique pour l'une peut découler d'une catastrophe pour les autres. En résumé, un gouvernement mondial efficace est nécessaire – un gouvernement fédéral qui autorisera les différences culturelles et qui (il faut l'espérer) garantira le respect des droits de l'homme.

Une telle chose est-elle possible ?

Peut-être.

Dans les pages qui précèdent, j'ai parlé de population mondiale et de taux de croissance en me basant sur des statistiques de 1970. Depuis cette date, il semblerait que le taux de croissance ait quelque peu diminué. Les gouvernements ont de plus en plus pris conscience de l'énorme danger que représente la surpopulation et se sont rendu compte que si l'on n'arrive pas à trouver une solution au problème, *aucun* autre ne pourra être résolu. Partout se développent des centres de planning familial, et la Chine (avec son milliard d'habitants, elle représente presque le quart de la population mondiale) encourage actuellement très fortement la famille à un seul enfant.

Il en résulte que le taux de croissance de la population mondiale est passé en 1970 de 2 % à 1,6 % (c'est une estimation) au début des années quatre-vingts. Il est bien certain que la population du globe est passée à 4,5 milliards, tant et si bien qu'une augmentation de 1,6 % équivaut à 72 millions d'habitants en plus chaque année – c'est-à-dire un tout petit peu plus que l'augmentation annuelle en 1970. Autrement dit, nous ne sommes pas encore allés assez loin, mais nous allons dans la bonne direction.

De plus, on assiste à une progression régulière du féminisme. Les femmes comprennent l'importance de partager les responsabilités dans tous les aspects de la vie quotidienne et sont de plus en plus prêtes à les assumer. L'importance de cet événement (hormis son aspect de pure équité) réside dans le fait que les femmes qui assument des responsabilités professionnelles trouvent de nouvelles possibilités de s'accomplir autrement que dans les rôles traditionnels de machines à faire des enfants et de femmes de ménage, et il est fort probable que le taux de natalité restera faible.

Mais il est également certain que le mouvement dans la direction du contrôle des naissances, aussi essentiel qu'il puisse paraître à n'importe quel esprit sensé, n'est pas sans opposants. Aux États-Unis, il existe un groupe très militant qui s'oppose non seulement à l'avortement, mais également à toutes les formes d'éducation sexuelle à l'école, et à la disponibilité de moyens anti-conceptionnels qui, pourtant, rendrait l'avortement inutile. Pour ces gens-là, la seule façon légitime d'abaisser le taux de natalité réside dans l'abstinence sexuelle totale ; seuls des fous pourraient penser que les gens vont accepter de pratiquer une telle méthode. Ce groupe s'est donné le nom de « Droit à la Vie », mais un nom plus approprié pour ceux qui refusent de reconnaître les dangers de la surpopulation serait « Droit à la Stupidité fatale ».

Et puis il y a eu les pays arabes producteurs de pétrole qui, en 1973, ont décidé un embargo temporaire sur les livraisons pour punir les pays occidentaux de ce qu'ils ressentaient comme un soutien inconditionnel à l'État d'Israël. Cette mesure, ainsi que les nombreuses années qui l'ont suivie, pendant lesquelles le prix du pétrole n'a cessé d'augmenter, ont convaincu les pays industrialisés de la nécessité de faire des économies

d'énergie. Si cette politique continue – et si l'on y associe une volonté résolue de remplacer les combustibles d'origine fossile, autant que cela se peut, par l'énergie solaire, la fusion nucléaire et les sources d'énergie renouvelables – nous aurons fait un pas immense pour la survie de notre espèce.

Nous nous faisons aussi de plus en plus de soucis au sujet de la qualité de notre environnement. Aux États-Unis, l'administration Reagan, qui est arrivée au pouvoir en 1981, a mis sur pied un certain nombre de programmes qui favorisent les grandes entreprises au mépris des grands idéaux humanitaires traditionnellement défendus depuis le New Deal de Franklin D. Roosevelt, cinquante ans auparavant. Pour cela, le gouvernement américain pensait pouvoir compter sur le soutien d'une majorité de ses électeurs. Mais quand on a vu arriver à la tête de l'Agence pour la protection de l'environnement des hommes qui pensaient que cela valait la peine d'empoisonner la majorité pour les profits d'une minorité, il y eut un tel tollé que l'Agence fut réorganisée et que l'administration fut bien obligée d'admettre qu'elle avait « mal interprété son mandat ».

Il ne faut pas non plus que nous sous-estimions les effets du progrès technologique, par exemple ceux de la révolution qui secoue les communications. La prolifération des satellites de communications permettra peut-être bientôt à chacun de se mettre en rapport avec n'importe quelle autre personne sur le globe. Les pays sous-developpés pourront ainsi se dispenser d'établir les anciens réseaux de communications qui nécessitaient un énorme investissement, pour entrer directement dans un monde dans lequel tout un chacun possède pour ainsi dire sa propre chaîne de télévision, capable d'envoyer et de recevoir des messages.

Le monde va devenir tellement plus petit qu'il ressemblera dans sa structure sociale à un véritable village. De fait, le terme *village global* a été utilisé pour décrire cette nouvelle situation. L'éducation va pénétrer dans chaque recoin de ce village grâce à l'ubiquité de la télévision. Dans chaque pays en voie de développement, la nouvelle génération pourra s'informer sur les dernières méthodes appliquées à l'agriculture, sur l'utilisation à bon escient des engrais et des pesticides, et sur les techniques de contraception.

Il se peut que pour la première fois dans l'histoire de la Terre, on assiste à une tendance vers la décentralisation. Avec la télévision et la possibilité qu'elle crée de transmettre une conférence d'hommes d'affaires, par exemple, ou un programme culturel aux quatre coins du monde, le besoin de tout concentrer dans une énorme masse en décomposition se fera moins sentir.

Les ordinateurs et les robots (dont je parlerai dans le prochain chapitre) auront peut-être aussi quelque effet salutaire.

Alors, qui sait ? Peut-être nous dirigeons-nous vers une catastrophe, mais nos efforts pour atteindre un éventuel salut ne seront peut-être pas vains.

VIVRE SOUS LES MERS

A supposer que nous réussissions à relever le défi, que la population se stabilise, et même qu'une lente décroissance ait lieu, qu'un gouvernement mondial efficace et responsable soit institué et qu'il autorise le droit à la différence, que la structure écologique soit protégée et que toute forme de vie sur Terre soit systématiquement préservée – que se passera-t-il alors ?

Tout d'abord, l'humanité continuera probablement à s'étendre. Après ses débuts d'hominidé primitif en Afrique orientale – et après avoir sans doute remporté un succès du même ordre que celui du gorille moderne – le prédécesseur de l'être humain s'est lentement déplacé jusqu'à ce que, il y a 15 000 ans, *Homo sapiens* colonise toute l'île mondiale (l'Asie, l'Afrique et l'Europe). Les hommes ont alors effectué leur progression vers les Amériques, l'Australie et même à travers les îles du Pacifique. Au début du XXe siècle, la population demeurait peu dense dans les régions particulièrement hostiles telles que le Sahara, le désert d'Arabie et le Groënland – mais à part l'Antarctique, aucune région du globe n'était entièrement inhabitée. Maintenant, même là-bas, il existe des stations scientifiques permanentes.

Jusqu'où l'homme va-t-il aller ?

La mer offre une réponse possible. La vie est née dans la mer, et c'est là qu'elle se porte le mieux en termes purement quantitatifs. Toutes les formes d'animaux terrestres, les insectes mis à part, se sont essayées au retour à la mer, vu ses ressources pratiquement inépuisables de nourriture et la relative uniformité de son environnement. Parmi les mammifères, des exemples tels la loutre, le phoque et la baleine montrent des stades progressifs de réadaptation à un environnement aquatique.

Pouvons-nous retourner vers la mer, non par le biais d'une lente modification évolutive de notre corps, mais grâce à des progrès technologiques plus rapides ? Protégé par l'épaisseur des parois métalliques des sous-marins et des bathyscaphes, l'être humain a conquis l'Océan jusque dans ses plus grandes profondeurs.

Pour la plongée ordinaire, point n'est besoin, bien sûr, d'équipements aussi sophistiqués. En 1943, l'océanographe français Jacques Yves Cousteau inventa le poumon aquatique. Cet appareil fournit de l'oxygène au plongeur grâce à une bonbonne d'air comprimé qu'il porte sur le dos et qui a permis l'essor actuel de la plongée sous-marine. On peut ainsi rester sous l'eau pendant plusieurs heures sans être enfermé dans des nacelles ou engoncé dans des scaphandres.

Cousteau a également été un pionnier des habitations sous-marines, dans lesquelles les hommes peuvent vivre pendant des périodes encore plus longues. En 1964, par exemple, deux hommes ont vécu pendant deux jours sous une tente remplie d'air à une profondeur de 432 pieds sous le niveau de la mer (l'un d'eux était Jon Lindbergh, le fils de l'aviateur). Des hommes sont restés plusieurs semaines à des profondeurs moindres.

Encore plus spectaculaires sont les expériences menées par le biologiste Johannes A. Kylstra, à l'université de Leyden (Pays-Bas), sur la respiration aquatique chez les mammifères. Les poumons et les branchies fonctionnent de façon similaire, après tout, sauf que ces dernières peuvent fonctionner à des niveaux plus bas d'oxygénation. Kylstra a utilisé une solution aqueuse dont la formule était suffisamment proche de celle du sang des mammifères pour éviter tout dommage aux poumons, puis l'a fortement oxygénée. Il s'est aperçu que des souris et des chiens étaient capables de « respirer » ce liquide pendant de longues périodes sans effets secondaires notables.

On a maintenu en vie des hamsters dans de l'eau ordinaire en les entourant d'un film très mince de silicone permettant à l'oxygène de passer de l'eau vers l'animal, et au gaz carbonique de passer de l'animal vers l'eau. La membrane, en l'occurence, faisait fonction de branchie. Ces

progrès, et d'autres encore dans le futur, permettront-ils à l'homme de rester sous l'eau indéfiniment, rendant ainsi toute la surface du globe – terre et mer – habitable ?

VIVRE DANS L'ESPACE

Et que dire de l'espace ? Resterons-nous sur notre planète, ou bien nous lancerons-nous à la découverte d'autres mondes ?

Après le lancement des premiers satellites en 1957, les rêves qu'avait fait l'homme de voyager dans l'espace, rêves entretenus jusque-là par les récits de science-fiction, ont commencé à prendre corps. Il n'a fallu que trois ans et demi après le lancement de *Spoutnik I* pour qu'un homme vole dans l'espace, et seulement huit ans après cet événement pour que des êtres humains foulent le sol de la Lune.

Mais le programme spatial a coûté très cher et a été vivement critiqué par bon nombre de chercheurs parce que trop orienté vers les relations publiques et pas assez vers des buts scientifiques, ou parce qu'il portait ombrage à d'autres programmes scientifiques plus urgents. Même le grand public a parfois émis des réserves, le trouvant trop cher, alors que d'autres priorités d'ordre sociologique assaillent le monde.

Et pourtant, le programme spatial continuera probablement son petit bonhomme de chemin, peut-être à vitesse réduite ; et si les hommes trouvent un jour le moyen de réduire leur consommation d'énergie et les sommes qu'ils consacrent traditionnellement à leur folie guerrière, peut-être pourra-t-il même accélérer. Il existe déjà des plans pour la construction de stations spatiales – en fait, d'énormes satellites en orbite plus ou moins permanente autour de la Terre et capables de recevoir des hommes et des femmes pendant de longues périodes – afin de pouvoir réaliser des observations et des expériences d'une grande portée scientifique. Des navettes, utilisables un grand nombre de fois, ont été mises au point, fonctionnent à merveille, et constituent l'étape préliminaire essentielle à la réalisation de ces objectifs.

On peut espérer que d'autres voyages vers la Lune auront lieu, qui permettront l'établissement là-bas de colonies plus ou moins permanentes, et l'exploitation des ressources lunaires ; à terme, ces colonies pourront devenir totalement autonomes.

Cependant, en 1974, le physicien américain Gerard Kitchen O'Neill a suggéré qu'il serait inutile d'établir une vaste colonie permanente sur la Lune, et qu'une petite station minière pourrait suffire. Bien que la vie ait commencé à la surface d'une planète, il n'était pas nécessaire qu'elle s'y confine. Il a proposé que de grands cylindres, sphères, ou anneaux soient placés sur orbite et qu'on leur imprime un mouvement de rotation assez rapide pour provoquer une force centrifuge qui maintiendrait les gens au sol sous l'effet d'une pesanteur artificielle.

De telles stations pourraient être construites en métal et en verre, et le sol serait recouvert de matériaux rocheux, le tout d'origine lunaire. On pourrait en aménager l'intérieur à l'image des villes terrestres et y faire vivre dix mille habitants ou plus, selon la taille de la station. L'orbite pourrait être à la position troyenne par rapport à la Terre et à la Lune (c'est-à-dire que la Terre, la Lune et la station se trouveraient aux trois coins d'un triangle équilatéral).

Il existe deux positions de ce genre, et plusieurs dizaines de stations pourraient s'agglutiner à chacune d'elles. Dans l'immédiat, ni les États-Unis ni l'Union Soviétique ne paraissent envisager de telles structures, mais le fougueux O'Neill pense que si l'humanité s'engageait à poursuivre un tel projet avec toute la vigueur qu'il mérite, il y aurait très vite plus d'êtres humains dans l'espace que sur Terre.

Les stations de O'Neill, au moins dans un premier temps, sont destinées à des orbites lunaires. Mais l'homme peut-il s'engager dans l'espace au-delà de la Lune ?

En théorie, aucune raison ne l'en empêche, mais les voyages vers l'autre monde le plus proche où il puisse se poser, Mars (bien que Vénus soit plus près de la Terre, son atmosphère est trop brûlante pour permettre des expéditions humaines), nécessiteront des vols non pas de plusieurs jours, comme c'est le cas pour la Lune, mais de plusieurs mois. Et pendant tous ces longs mois, l'homme devra emmener avec lui une atmosphère dans laquelle il puisse vivre.

Les êtres humains ont déjà une certaine expérience en la matière, puisqu'ils sont descendus au fond des océans dans des sous-marins et des engins comme le bathyscaphe. De même, ils partiront à la conquête de l'espace dans des bulles d'air, entourées d'une solide coque métallique, emportant avec eux la nourriture, l'eau et tout ce dont ils auront besoin au cours du voyage. Mais c'est le lancement qui pose le plus de problèmes, en particulier celui de vaincre la pesanteur. Une grande proportion du poids et du volume d'un vaisseau spatial doit être consacrée au groupe propulseur et au carburant, et la charge utile réservée à l'équipage et à ses provisions sera tout d'abord très réduite.

Il faudra que les vivres soient très compactes ! pas de place pour les aliments indigestes. La nourriture artificielle et condensée sera composée de lactose, d'une huile végétale légère, d'un mélange approprié associant acides aminés, vitamines, minéraux, avec un soupçon d'assaisonnement, le tout dans une petite boîte fabriquée en hydrates de carbone comestibles. Une boîte contenant 180 grammes de nourriture solide suffirait à un repas. Trois boîtes fourniraient 3 000 calories. Il faudrait ajouter à cela 1 gramme d'eau par calorie (entre 2,5 et 3 litres d'eau par jour et par personne) ; une partie de cette eau pourrait être mélangée à la nourriture pour rendre cette dernière plus agréable au goût, nécessitant une boîte plus grande. De plus, le vaisseau spatial devrait emporter l'oxygène nécessaire pour respirer, c'est-à-dire environ 1 litre (1 150 grammes) d'oxygène sous forme liquide par jour et par personne.

Les besoins quotidiens pour une personne s'élèveraient ainsi à 540 grammes de nourriture solide, 2 700 grammes d'eau, et 1 150 grammes d'oxygène. Total : 4 390 grammes. Imaginez donc un voyage vers la Lune, qui va prendre une semaine dans chaque sens, en prévoyant deux jours à la surface de la Lune pour son exploration. Chaque passager à bord du vaisseau aura besoin d'environ 70 kilos de nourriture, d'eau et d'oxygène. La technologie actuelle nous le permet sans grandes difficultés.

Mais les exigences d'une expédition vers Mars sont bien supérieures. Une telle expédition prendrait plus de deux ans et demi, en laissant suffisamment de temps sur Mars pour attendre une phase favorable des positions respectives des planètes autorisant le retour. Sur la base des chiffres que je viens d'évoquer, un tel vol nécessiterait 5 tonnes de nourriture, d'eau et d'oxygène par personne. Transporter un tel poids de provisions dans un vaisseau spatial est actuellement impensable d'un point de vue technologique.

La seule solution raisonnable pour effectuer un long périple dans l'espace consiste à rendre le vaisseau autonome, telle la Terre, elle-même énorme « vaisseau » qui voyage dans l'espace. La nourriture, l'eau et l'air d'origine devront être réutilisés indéfiniment par le recyclage des déchets.

De tels *systèmes fermés* existent déjà en théorie. Il n'est pas très agréable de parler de recyclage des déchets, mais c'est, après tout, le processus qui maintient la vie sur Terre. Des filtres chimiques embarqués seront capables de recueillir le gaz carbonique et la vapeur d'eau exhalés par les membres de l'équipage ; l'urée, le sel et l'eau seront récupérés par distillation et par d'autres procédés à partir de l'urine et des matières fécales ; les résidus secs de ces dernières seront stérilisés aux ultraviolets et, de même que le gaz carbonique et l'eau, serviront à nourrir des algues. Grâce à la photosynthèse, les algues convertiront le gaz carbonique et les substances azotées des matières fécales en aliments organiques, plus de l'oxygène, pour l'équipage. Seule une source d'énergie extérieure au système sera nécessaire pour alimenter les divers processus de recyclage, y compris la photosynthèse ; le soleil fera parfaitement l'affaire.

On estime qu'environ 110 kilos d'algues par voyageur seront suffisants pour subvenir indéfiniment aux besoins en nourriture et en oxygène de l'équipage. Si l'on ajoute le poids de l'équipement nécessaire à tous les traitements des déchets, le poids total du matériel requis, par personne, serait d'environ 160 kilos, mais certainement pas plus de 500 kilos. On a également étudié des systèmes basés sur les bactéries qui utilisent l'hydrogène. Ces systèmes n'ont pas besoin de lumière, mais seulement d'hydrogène, que l'on peut facilement obtenir par électrolyse de l'eau. Les résultats montrent que ces systèmes ont un rendement bien supérieur à celui des organismes qui ont recours à la photosynthèse.

Mais à part les problèmes de matériel, il y a également celui de l'absence de pesanteur auquel seront soumis les futurs voyageurs de l'espace pendant de longues périodes. Des astronautes ont vécu six mois en état d'apesanteur continu sans dommage permanent, mais ils ont assez souffert de troubles mineurs pour que l'on s'inquiète de ses effets prolongés. Heureusement, il existe des moyens de s'en affranchir. Une lente rotation du véhicule spatial, par exemple, produira la sensation de pesanteur par l'effet de la force centrifuge.

Plus sérieux et moins facilement écartés sont les dangers associés aux brusques accélérations et décélérations que les astronautes doivent subir au lancement et à l'atterrissage de leur engin.

La force normale de la gravité à la surface de la Terre est de 1 *g*. L'état d'apesanteur est caractérisé par 0 *g*. Une accélération (ou une décélération) qui double le poids du corps est de 2 *g*, une force qui le triple de 3 *g*, etc.

La position du corps, pendant l'accélération, est un facteur primordial. Si l'accélération a lieu « tête la première » (ou s'il y a décélération « pieds en avant »), le sang reflue de la tête. Si l'accélération est assez forte (disons 6 *g* pendant 5 secondes), il se produit un *trou noir.* D'un autre côté, si l'accélération a lieu « pieds en avant » (ou *accélération négative,* au contraire de l'accélération *positive,* tête la première), le sang afflue vers la tête. Ce dernier phénomène est beaucoup plus dangereux, car l'augmentation de la pression risque de faire éclater les vaisseaux sanguins des yeux et du cerveau. Les spécialistes l'appellent le *voile rouge*. Une accélération de 2,5 *g* pendant 10 secondes est suffisante pour endommager certains vaisseaux.

L'accélération transversale est la plus facilement tolérée – c'est-à-dire lorsque la force est appliquée à angle droit par rapport à l'axe longitudinal du corps, telle qu'en position assise par exemple. Certains pilotes d'essais ont subi des accélérations transversales aussi élevées que 10 *g* pendant plus de 2 minutes dans une centrifugeuse sans perdre conscience.

Pendant des laps de temps plus courts, la tolérance est encore plus grande. Des records étonnants de décélération ont été battus par un colonel, John Paul Stapp, ainsi que d'autres volontaires sur la piste de la base aérienne américaine d'Holloman, dans le Nouveau-Mexique. Au cours d'une célèbre expérience qui a eu lieu le 10 décembre 1954, Stapp a subi une décélération de 25 *g* pendant environ 1 seconde. Le chariot qui l'emportait sur des rails à près de 1 000 km/h a freiné pour s'arrêter complètement en moins d'une seconde et demi, comme si une voiture percutait un mur de briques à 200 km/h ! Bien sûr, Stapp était ficelé dans son chariot afin de limiter les risques au maximum ! Il s'en est tiré avec quelques contusions, des ampoules, et un traumatisme oculaire qui lui a valu deux superbes yeux au beurre noir.

Au décollage, un astronaute doit absorber (pendant un court instant) quelque 6,5 *g* et au retour, un maximum de 11 *g*.

Des dispositifs tels que des couchettes épousant la forme du corps, divers types de harnais, ou bien encore l'immersion dans une capsule ou une combinaison spatiale remplie d'eau, offriront une marge de sécurité suffisante contre les *g* très élevés.

Les dangers posés par les radiations sont également à l'étude, tout comme l'ennui provoqué par de longues périodes de solitude, l'étrange expérience de l'espace sans bruit et où la nuit ne tombe jamais, ainsi que d'autres conditions inhabituelles auxquelles les voyageurs de l'espace seront confrontés. Mais l'un dans l'autre, ceux qui se préparent à ces longs voyages loin de la planète natale ne voient aucun obstacle majeur à cette grande première pour l'humanité.

Les difficultés psychologiques posées par ces longues expéditions ne s'avéreront pas aussi sérieuses que nous l'imaginons si nous abandonnons l'idée préconçue de l'astronaute en tant que terrien. Il est évident que pour ce dernier, il existe des différences énormes entre la vie à la surface d'une immense planète et la vie dans un petit vaisseau spatial.

Mais qu'en sera-t-il quand ces explorateurs d'un type nouveau seront des habitants des colonies du type envisagé par O'Neill ? Ces hommes et ces femmes seront habitués à un environnement confiné, au recyclage de leur nourriture, de leur boisson et de leur air, aux variations des forces gravitationnelles ; en bref, ils seront habitués à vivre dans un environnement spatial. Un vaisseau spatial ne sera qu'une version miniaturisée de la colonie dans laquelle ils auront peut-être passé toute leur vie.

Il est très possible que les colons de l'espace deviendront les fers de lance de l'exploration humaine au XXIe siècle et après. Ce seront eux, peut-être, qui atteindront les astéroïdes. Là, des activités minières fourniront de nouvelles ressources à une humanité en expansion, et plusieurs astéroïdes pourront être creusés pour abriter de nouvelles colonies naturelles, dont de nombreuses seront beaucoup plus peuplées que celles permises par le système Terre-Lune.

A partir des astéroïdes, les hommes pourront alors commencer à explorer les vastes espaces aux confins du système solaire... et au-delà, les étoiles.

Chapitre 17

L'esprit

Le système nerveux

Physiquement parlant, nous autres êtres humains sommes des spécimens plutôt ordinaires comparés à d'autres organismes. Nous aurions beaucoup de mal à concurrencer n'importe quel animal de notre taille dans une quelconque épreuve de force. Nous marchons sans grâce, par rapport au chat par exemple ; nous sommes incapables de suivre un chien ou un cerf à la course ; pour ce qui est de la vision, de l'ouïe ou de l'odorat, d'autres animaux nous surpassent et de loin. Notre squelette n'est pas vraiment adapté à la station debout : nous sommes vraisemblablement les seuls animaux à souffrir du dos tout en menant une activité normale. Quand, du point de vue de l'évolution, nous pensons à la perfection qui caractérise certains animaux – la splendide efficacité du poisson dans son milieu aquatique ou de l'oiseau dans les airs, la grande fécondité et la faculté d'adaptation des insectes, la simplicité et l'efficacité simplement géniales des virus – l'être humain apparaît vraiment comme une créature maladroite et bien mal conçue. En tant qu'organismes vivants, nous aurions toutes les peines du monde à concurrencer n'importe quelle créature dans son environnement. Mais nous sommes malgré tout parvenus à dominer la planète grâce à une spécialisation tout à fait particulière d'un organe – le cerveau.

LES CELLULES NERVEUSES

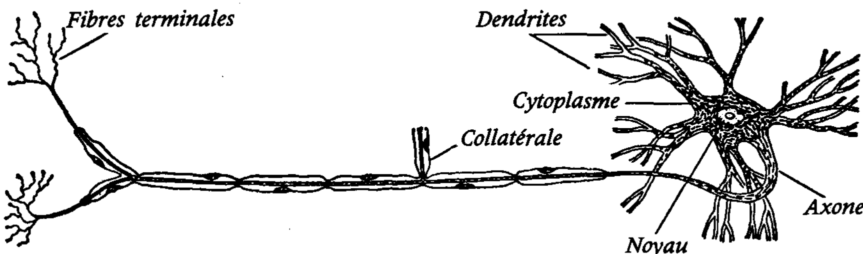
Une cellule est sensible aux changements qui interviennent dans son environnement (*stimuli*) et réagit en conséquence (*réponses*). C'est ainsi

qu'un protozoaire nage en direction d'une goutte de solution sucrée déposée dans l'eau non loin de lui, mais fuit une goutte d'acide. Il est bien évident que ce type de réponse directe et automatique convient parfaitement à une cellule isolée, mais créerait un véritable chaos pour un ensemble de cellules. Tout organisme composé d'un certain nombre de cellules doit posséder un système qui assure la coordination de leurs réponses. Sans un tel système, cet organisme ressemblerait à une ville dans laquelle les habitants ne pourraient communiquer entre eux et où chacun tirerait à hue et à dia. Même les cœlenterata, les animaux pluricellulaires les plus primitifs, possèdent les rudiments d'un système nerveux. On peut y voir les premières cellules nerveuses (les *neurones*) – des cellules spécialisées dont les fibres se ramifient à partir du corps de la cellule pour former de petites branches extrêmement fines (figure 17.1).

Le fonctionnement des neurones est d'une telle complexité que, déjà à ce niveau pourtant relativement simple, l'explication exacte de tous les phénomènes qui s'y déroulent dépasse actuellement notre entendement. Il se passe quelque chose (nous ne savons pas encore très bien quoi) lorsqu'un changement se produit dans l'environnement de la cellule. Ce changement peut avoir la forme d'une modification dans la concentration d'une certaine substance, dans la quantité de lumière, dans les mouvements de l'eau, ou bien il peut s'agir du contact direct avec un objet. Quel que soit le stimulus, il induit un petit *influx nerveux*, un courant électrique qui se propage rapidement le long de la fibre. Lorsqu'il atteint l'extrémité de cette fibre, cet influx franchit un passage étroit (appelé *synapse*) vers la cellule nerveuse voisine ; et c'est ainsi qu'il se transmet de cellule à cellule. Dans un système nerveux bien développé, un neurone peut être relié à ses voisins par des milliers de synapses. Chez les cœlenterata, comme la méduse par exemple, l'influx se propage dans tout l'organisme. La méduse répond en contractant tout ou partie de son corps. Si le stimulus provient d'une particule d'aliment, l'organisme l'absorbe en contractant ses tentacules.

Tout cela, bien entendu, est entièrement automatique, mais comme la méduse y trouve son compte, nous sommes tentés de voir, dans la façon dont se comporte cet organisme, un but, une finalité. Très naturellement, l'homme, en tant que créature douée de raison et poursuivant des buts dans la vie, prête ces attributs même aux animaux inférieurs. Les savants voient là une attitude *téléologique*, et font tout leur possible pour éviter ce travers dans leur mode de pensée et dans leur discours. Mais lorsqu'on

Figure 17.1. Une cellule nerveuse.



en vient à décrire les résultats de l'évolution, il est si facile de parler en termes de développement finaliste de l'organisme vers une plus grande efficacité, qu'à part certains puristes des plus fanatiques, mêmes les scientifiques tombent parfois dans ce travers (vous vous êtes certainement rendu compte, chers lecteurs, que j'ai moi-même bien souvent péché). Essayons, cependant, d'éviter cet écueil lorsque nous étudions le développement du système nerveux et du cerveau. Ce n'est pas la nature qui a conçu le cerveau ; ce dernier est le résultat d'une longue série d'accidents, pour ainsi dire, qui, au cours de l'évolution, ont produit certaines caractéristiques qui se sont avérées avantageuses pour les organismes qui les possédaient. Dans la lutte pour la survie, un animal plus sensible aux modifications de son environnement que ses concurrents et plus rapide qu'eux dans ses réponses se trouvait favorisé par le processus de la sélection naturelle. Si, par exemple, un animal possédait des taches particulièrement sensibles à la lumière, cet avantage revêtirait une telle importance que l'évolution de taches oculaires, et plus tard d'yeux, ne manquerait pas de s'ensuivre.

Des ensembles spécialisés de cellules, l'équivalent d'*organes sensoriels* rudimentaires, apparaissent chez les plathelminthes, c'est-à-dire les vers plats. On note également chez ces derniers l'esquisse d'un système nerveux qui n'envoie pas des influx nerveux dans le corps entier au hasard, mais les dirige vers certains points de réponse critiques. L'organe qui effectue ce travail est un cordon nerveux central. Les vers plats sont les premiers animaux chez qui l'on constate l'existence d'un *système nerveux central.*

Et ce n'est pas tout. Les organes sensoriels du ver plat sont localisés dans l'extrémité qui lui sert de tête, la première partie de son corps à rencontrer des obstacles quand il se meut, et le cordon nerveux est donc particulièrement développé dans cette région antérieure. Il s'agit là d'un embryon de cerveau.

Petit à petit, les embranchements plus complexes se dotent de certaines variantes. Le nombre des terminaisons nerveuses augmente, et ces dernières deviennent plus sensibles. Le cordon nerveux et ses ramifications deviennent plus compliqués, en créant un vaste système de cellules nerveuses *afférentes* (« transmettant vers »), qui transfèrent les messages vers le cordon, et de fibres *efférentes* (« transmettant de »), qui les transmettent vers les organes chargés de la réponse. Le nœud de cellules nerveuses à la croisée des chemins, dans la tête, devient de plus en plus complexe. Les fibres nerveuses évoluent et transmettent l'influx nerveux de plus en plus vite. Chez le calmar, le plus développé de tous les animaux non segmentés, cette transmission plus rapide s'effectue grâce à un épaississement de la fibre nerveuse. Chez les animaux segmentés, la fibre se dote d'une gaine de matériau gras (la *myéline*), qui est encore plus efficace pour la transmission de l'influx nerveux. Chez l'être humain, certaines fibres nerveuses peuvent transmettre l'influx à une vitesse de 100 m/s (plus de 350 km/h), comparée à seulement environ 160 m/h chez certains invertébrés.

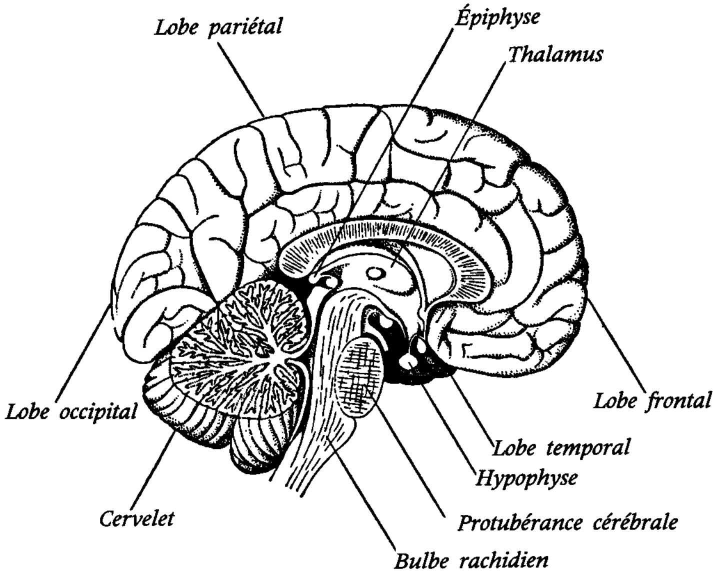
Les cordés introduisent un changement radical dans l'emplacement du cordon nerveux. Le tronc nerveux principal (aussi connu sous le vocable de *cordon médullaire*) se trouve le long du dos et non plus le long du ventre, comme chez les animaux inférieurs. A première vue, on pourrait croire qu'il s'agit là d'un pas en arrière – le cordon est placé à un endroit plus exposé. Mais les vertébrés sont dotés d'une colonne vertébrale osseuse qui le protège efficacement. Bien que sa fonction première soit de protéger

le cordon médullaire, la colonne vertébrale présente des avantages remarquables, car elle sert de solive à laquelle les cordés peuvent suspendre volume et poids. Des côtes se forment à partir de cette colonne et entourent la poitrine, ainsi que des mâchoires munies de dents pour mâcher, et des os longs qui constituent des membres.

LE DÉVELOPPEMENT DU CERVEAU

Le cerveau des cordés se développe à partir de trois structures qui existent déjà à l'état embryonnaire chez les vertébrés les plus primitifs. Ces structures, au départ de simples renflements de tissu nerveux, forment le *prosencéphale* (« cerveau antérieur »), le *mésencéphale* (« cerveau moyen ») et le *rhombencéphale* (« cerveau postérieur »), une distinction que l'on doit à l'anatomiste grec Erasistrate, vers 280 av. J.-C. Le cordon médullaire, à son extrémité supérieure, s'élargit régulièrement pour former la section postérieure du cerveau, appelée *bulbe rachidien*. Sur la partie antérieure de cette section, à part chez les cordés les plus primitifs, se trouve une sorte de bosse qu'on appelle le *cerebellum* (le cervelet, ou « petit cerveau »), et à l'avant de ce dernier, le cerveau moyen. Chez les vertébrés inférieurs, le cerveau moyen est le siège de l'odorat et du goût, et contient les *bulbes olfactifs*. Le cerveau antérieur, de l'avant vers l'arrière, est divisé en trois parties : la section des bulbes olfactifs, le cerveau, et le *thalamus*, dont la partie inférieure s'appelle l'*hypothalamus*. Chez les humains, en tout cas, le cerveau constitue la partie la plus développée et la plus importante de tout l'organe. En retirant le cerveau de certains animaux, l'anatomiste français Marie Jean Pierre Flourens a pu montrer en 1824 que c'est effectivement le cerveau qui est le siège de la pensée et de la volonté. (La figure 17.2 donne un schéma du cerveau humain.)

Figure 17.2. Le cerveau humain.



De plus, c'est le haut du cerveau, le *cortex cérébral*, qui joue le rôle le plus important. Chez les poissons et les amphibiens, il ne s'agit que d'une enveloppe lisse (le *pallium*, ou « manteau »). Chez les reptiles, on voit apparaître un morceau de tissu nerveux, le *néopallium* (le « nouveau manteau »). C'est un précurseur d'une grande importance, car c'est lui qui va se charger de gérer la vision et bien des sensations. Chez les reptiles, le centre des messages visuels est déjà partiellement passé du cerveau moyen au cerveau antérieur ; chez les oiseaux, le saut est accompli. Chez les premiers mammifères, c'est le néopallium qui commence à prendre la relève. Il s'étend pratiquement sur toute la surface du cerveau. C'est tout d'abord un revêtement lisse, mais au cours de son développement chez les mammifères supérieurs, il commence à former des plis, des *circonvolutions*. Ces plissements sont responsables de la complexité et du volume du cerveau chez les mammifères supérieurs, particulièrement chez *Homo sapiens*.

Plus les espèces se développent, plus on est frappé du rôle croissant qu'assume le cerveau. Dans le cas des primates, chez qui la vue prend de l'importance au détriment de l'odorat, les lobes olfactifs du cerveau antérieur se réduisent à leur plus simple expression. En même temps, le cerveau se développe et recouvre le thalamus et le cervelet.

Même les fossiles d'humanoïdes les plus anciens possédaient un cerveau considérablement plus volumineux que celui des singes les plus développés. Alors que le cerveau du chimpanzé ou de l'orang-outan pèse moins de 400 grammes, que celui du gorille, bien que plus gros que celui de l'homme, atteint en moyenne environ 540 grammes, le cerveau de *Pithecanthropus* pesait apparemment entre 850 grammes et 1 kilogramme. Et il s'agit là d'hominidés possédant de *petits cerveaux*. Le cerveau de l'homme de Rhodésie pesait environ 1,3 kilogramme ; celui de l'homme de Néanderthal et de l'*Homo sapiens* moderne se situe aux alentours de 1,5 kilogramme. On peut expliquer la supériorité mentale de l'homme actuel sur l'homme de Néanderthal par le fait qu'une grande proportion du cerveau humain est concentrée dans les régions antérieures, celles qui, semble-t-il, contrôlent les fonctions mentales supérieures. Néanderthal avait le front bas et son cerveau se présentait sous la forme d'une protubérance postérieure ; au contraire, l'homme actuel possède un front haut et son cerveau fait saillie vers l'avant.

La taille du cerveau des hominidés a environ triplé au cours des trois derniers millions d'années – une augmentation rapide à l'échelle des modifications engendrées par l'évolution. Mais pourquoi cette augmentation ? Et pourquoi chez les hominidés ?

Une raison probable en est que, comme nous le savons maintenant, les hominidés dotés de petits cerveaux marchaient debout, comme l'homme actuel. Une longue période de station debout a précédé le développement du cerveau. De cette attitude verticale découlent en effet deux conséquences importantes : tout d'abord, les yeux se retrouvent plus haut au-dessus du sol et fournissent au cerveau une quantité accrue d'informations ; deuxièmement, les membres antérieurs se retrouvent libres *de façon permanente*, ce qui leur permet de toucher et de manipuler l'environnement. L'afflux de perceptions sensorielles, la vue d'objets à une certaine distance et le toucher d'objets proches nécessitent un cerveau développé, capable de traiter ces nouvelles données, et tout individu doté d'un cerveau plus efficace (soit parce qu'il est plus gros, soit parce qu'il est mieux

organisé) va se trouver par là même avantagé par rapport aux autres. Les processus de l'évolution produiront inévitablement des hominidés dotés de gros cerveaux. (Ce que nous savons maintenant nous permet en tout cas de l'affirmer.)

LE CERVEAU HUMAIN

Le cerveau de l'homme moderne représente 1/50 du poids total du corps humain. Chaque gramme du cerveau assume donc la charge de 50 grammes du corps. En comparaison, le cerveau du chimpanzé représente environ 1/150 du poids de son corps, et celui du gorille à peu près 1/500. Chez certains primates plus petits, il existe des rapports cerveau/corps plus élevés que chez l'homme. C'est d'ailleurs aussi le cas chez l'oiseau-mouche. Un ouistiti, par exemple, peut avoir un cerveau qui pèse 1/18 du poids de son corps. Mais dans ce cas, la masse du cerveau est trop petite en valeur absolue pour être capable d'offrir la complexité nécessaire à une intelligence de l'ordre de celle de l'homme. En résumé, ce qui est nécessaire, et l'homme a la chance d'en être doté, c'est un cerveau qui soit à la fois gros en valeur absolue et par rapport à la masse corporelle.

Cela apparaît évident lorsqu'on constate que deux types de mammifères possèdent un cerveau beaucoup plus gros que celui de l'homme, ce qui pourtant ne leur procure pas une intelligence surhumaine. Les plus gros éléphants peuvent posséder des cerveaux de plus de six kilos et les plus grosses baleines des cerveaux qui pèsent jusqu'à neuf kilos. La masse des corps que ces cerveaux doivent gérer est, par contre, énorme. Le cerveau de l'éléphant, malgré son poids, ne représente que 1/1 000 du poids de son corps, et le cerveau d'une grosse baleine 1/10 000 seulement.

Mais les humains ont deux rivaux possibles, et deux seulement. Le dauphin et le marsouin, de petits membres de la famille de la baleine, sont en effet des concurrents sérieux. Certains d'entre eux ne pèsent pas plus qu'un homme, et sont pourtant dotés d'un cerveau plus gros (pouvant atteindre 1,7 kilo) qui possède des circonvolutions plus nombreuses.

Il ne faut pas en déduire automatiquement que le dauphin est plus intelligent que nous, car il ne faut pas oublier la question de l'organisation interne du cerveau. Le cerveau du dauphin (comme celui de l'homme de Néanderthal) est peut-être plus orienté vers des tâches que l'on pourrait qualifier de subalternes.

La seule façon de le savoir, c'est d'essayer de mesurer l'intelligence du dauphin en l'étudiant scientifiquement. Des chercheurs comme John C. Lilly sont convaincus que l'intelligence du dauphin est effectivement comparable à la nôtre, que les dauphins et les marsouins possèdent un langage aussi compliqué que le nôtre, et qu'il est probablement possible d'établir une forme de communication entre leur espèce et la nôtre.

Même si c'est le cas, il ne fait aucun doute que les dauphins, aussi intelligents soient-ils, ont perdu l'occasion de traduire cette intelligence en un quelconque contrôle de leur environnement lorsqu'ils se sont réadaptés à la vie aquatique. Il est impossible d'utiliser le feu sous l'eau, et c'est cette découverte de l'utilisation du feu qui a démarqué les hominidés de tous les autres organismes. Plus important encore, une locomotion rapide dans un élément aussi visqueux que l'eau requiert une forme parfaitement hydrodynamique. Le dauphin s'est donc trouvé dans

l'impossibilité de se doter de l'équivalent du bras ou de la main de l'homme, grâce auxquels l'environnement peut être minutieusement examiné et manipulé.

Autre point intéressant, les hommes ont rattrapé et surpassé les cétacés. Alors que les hominidés possédaient encore un petit cerveau, celui des dauphins était déjà gros, et pourtant ces derniers n'ont pu arrêter l'inexorable progression des hommes. Il nous semblerait aujourd'hui inconcevable de laisser jouer les mécanismes de l'évolution et d'assister les bras croisés au développement d'un rat, ou même d'un chien, qui posséderait un cerveau plus gros que le nôtre, et menacerait ainsi notre position sur la Terre. Mais les dauphins, pris au piège des océans, n'ont rien pu faire pour empêcher le développement des hominidés, à tel point que nous pouvons aujourd'hui supprimer tous les cétacés presque sans effort, si nous en avions bien sûr l'intention. (Nous devons d'ailleurs tirer une certaine fierté du fait qu'autant d'entre nous ne le souhaitent pas, et fassent même tout pour l'éviter.)

Il est possible, d'un point de vue philosophique que nous ne comprenons pas encore, que les dauphins nous surpassent dans une forme quelconque d'intelligence, mais pour ce qui est du contrôle effectif de l'environnement et du développement de la technologie, *Homo sapiens* n'a pas d'égal sur Terre dans le temps présent ou, autant que l'on sache, n'en a pas eu dans le passé. (Il va sans dire que l'activité intelligente et technologique de l'homme, au moins telle qu'elle s'est déroulée jusqu'à présent, n'a pas toujours joué nécessairement en faveur de la planète – ni même d'ailleurs en sa faveur à lui.)

LES TESTS D'INTELLIGENCE

J'évoquais un peu plus haut les difficultés rencontrées pour déterminer précisément le niveau d'intelligence d'une espèce comme celle des dauphins ; profitons de cette occasion pour souligner qu'il n'existe aucune méthode pour mesurer de manière satisfaisante l'intelligence des membres de notre propre espèce.

En 1904, les psychologues français Alfred Binet et Théodore Simon ont mis au point un système permettant de mesurer l'intelligence au moyen de réponses données à certaines questions judicieusement choisies. Ces *tests d'intelligence* ont donné naissance à l'expression *quotient intellectuel* (ou QI), qui représente le rapport entre l'âge mental, mesuré par le test, et l'âge chronologique ; ce rapport est ensuite multiplié par cent pour se débarrasser des décimales. Ce sont principalement les travaux du psychologue américain Lewis Madison Terman qui ont popularisé le QI.

Le problème est le suivant : il n'existe pas de test qui ne soit relié à un contexte culturel donné. Des questions simples sur les charrues colleraient sans doute un jeune citadin intelligent, alors que des questions simples sur les escaliers mécaniques jetteraient dans un abîme de perplexité un enfant de la campagne tout aussi intelligent. Les deux types de questions ne manqueraient pas de laisser perplexe un petit aborigène australien tout aussi intelligent, qui, lui, aurait certainement des questions sur les boomerangs qui nous donneraient pas mal de fil à retordre.

Qui plus est, les gens ne peuvent s'empêcher d'avoir certains préjugés sur qui est intelligent et qui ne l'est pas. Un chercheur est quasi obligé

de constater une intelligence supérieure chez les sujets qui partagent avec lui le même milieu culturel. Dans son livre, *La Mal-mesure de l'homme* (1983), Stephen Jay Gould décrit en détail comment les mesures de QI ont été mises au service d'un racisme inconscient et banalisé.

L'exemple le plus récent et le plus flagrant en est celui du psychologue anglais Cyril Lodowic Burt, formé à Oxford et enseignant à la fois à Oxford et à Cambridge. Il a étudié le QI de certains enfants et l'a corrélé à la profession des parents : cadres de direction, cadres moyens, employés, ouvriers qualifiés, ouvriers semi-qualifiés, ouvriers non qualifiés.

Il a bien entendu trouvé que le QI était parfaitement corrélé avec les professions. Plus le parent était au bas de l'échelle sociale, plus le QI de l'enfant était bas lui aussi. Autrement dit, les gens appartenaient bien à leur milieu, et ceux qui étaient en haut de l'échelle le méritaient bien.

De plus, Burt a trouvé que les hommes avaient des QI plus élévés que les femmes, les Anglais plus que les Irlandais, les goys plus que les juifs, etc. Il a testé de vrais jumeaux séparés peu de temps après leur naissance pour trouver que leurs QI étaient néanmoins très proches – démontrant ainsi, une fois de plus, la primauté de l'hérédité sur l'environnement.

Avant sa mort en 1971, Burt fut comblé d'honneurs. Après sa mort, cependant, on découvrit que, sans l'ombre d'un doute, il avait manipulé ses résultats.

Point n'est besoin de rentrer dans ses motivations psychologiques. Je ne dirai qu'une chose : certains chercheurs sont tellement aveuglés par leurs préjugés qu'ils sont totalement incapables d'admettre des résultats qui contrediraient leurs hypothèses. Le domaine des tests d'intelligence est tellement chargé de facteurs subjectifs et d'amour-propre que tout résultat doit être considéré avec prudence.

Un autre test bien connu se propose de révéler des facettes de notre esprit encore plus insaisissables que l'intelligence. Il consiste en divers motifs sous forme de taches d'encre, préparées à l'origine par un médecin suisse, Hermann Rorschach, entre 1911 et 1921. On demande aux sujets quelles images ces taches d'encre évoquent pour eux ; à partir du type d'images qu'une personne construit dans le *test de Rorschach,* le spécialiste tire certaines conclusions sur sa personnalité. Même dans le meilleur des cas, on peut se demander à quel point les « conclusions » tirées sont vraiment « concluantes ».

LA SPÉCIALISATION DES FONCTIONS

Aussi étrange que cela puisse paraître, la plupart des anciens philosophes sont passés complètement à côté des phénomènes qui se produisent sous notre crâne. Aristote pensait que le cerveau n'était, pour ainsi dire, qu'une espèce de machine à réguler la température, servant à refroidir le sang quand celui-ci s'échauffait. A la génération suivante, Hérophile, qui habitait Alexandrie, identifia correctement le cerveau comme le siège de l'intelligence, mais, comme d'habitude, les erreurs d'Aristote eurent plus de poids que les vérités des autres.

Les penseurs de l'Antiquité et du Moyen Age ont donc eu tendance à placer le centre des émotions et de la personnalité dans des organes tels que le cœur, le foie, et la rate (comme en témoignent les expressions « avoir le cœur brisé », ou « se dilater la rate »).

C'est le médecin et anatomiste anglais Thomas Willis qui, au XVIIe siècle, a été le premier investigateur moderne du cerveau ; on lui doit la découverte des nerfs qui remontent jusqu'au cerveau. Plus tard un anatomiste français, Félix Vicq d'Azyr, et d'autres établirent un schéma général de l'anatomie du cerveau. Mais c'est au physiologiste suisse Albrecht von Haller, au XVIIIe siècle, qu'on doit la première découverte, cruciale, sur le fonctionnement du système nerveux.

Von Haller découvrit qu'il pouvait beaucoup plus facilement contracter un muscle en stimulant un nerf qu'en stimulant le muscle lui-même. De plus, cette contraction était involontaire, et il pouvait l'obtenir en stimulant un nerf même après la mort de l'organisme. Von Haller montra également que les nerfs transmettent les sensations. Lorsqu'il coupait les nerfs de certains tissus spécifiques, ces tissus ne pouvaient plus réagir. Le physiologiste en conclut que le cerveau reçoit les sensations par l'intermédiaire des nerfs, puis envoie, toujours par l'intermédiaire des nerfs, des messages qui provoquent une réponse, telle la contraction musculaire. Il émit l'hypothèse que tous les nerfs forment une jonction au centre du cerveau.

En 1811, le médecin autrichien Franz Joseph Gall étudia plus particulièrement la *substance grise* à la surface du cerveau (par opposition à la *substance blanche* qui consiste essentiellement en fibres provenant des cellules nerveuses, de couleur blanche à cause de leur enveloppe graisseuse). Gall suggéra que les nerfs ne se réunissent pas au centre du cerveau comme le pensait von Haller, mais que chacun d'eux rejoint une région bien définie de la substance grise, qu'il considérait comme le siège de la coordination du cerveau. Gall pensait que les différentes parties du cortex cérébral sont chargées de collecter les sensations des différentes parties du corps, et d'envoyer, en réponse, des messages à certaines régions spécifiques.

Si une zone définie du cortex est responsable d'une propriété bien particulière de l'esprit, quoi de plus naturel que de supposer que le degré de développement de cette partie indiquera le caractère ou la mentalité d'un individu ? En cherchant des bosses sur le crâne de quelqu'un, on pourrait ainsi se rendre compte si telle ou telle zone est plus développée, et dire s'il est particulièrement généreux, particulièrement dépravé, ou que sais-je encore. En partant de ce genre de raisonnement, certains disciples de Gall ont fondé une pseudo-science, la *phrénologie,* qui a connu un certain succès au XIXe siècle et a même des adeptes de nos jours. (Aussi étrange que cela puisse paraître, et bien que Gall et ses partisans aient vu dans un front haut et un crâne bien rond des signes d'intelligence supérieure – une idée encore très répandue aujourd'hui – Gall avait un cerveau particulièrement petit, environ 15 % plus petit que la moyenne.)

Mais le fait que la phrénologie, telle qu'elle s'est développée par l'entremise de charlatans, soit un tissu d'absurdités, ne signifie pas que l'idée originelle de Gall sur la spécialisation des fonctions dans des zones bien particulières du cortex cérébral était fausse. Bien avant de tenter des explorations spécifiques du cerveau, on s'était aperçu qu'un traumatisme dans une zone définie du cerveau causait une incapacité bien spécifique. En 1861, le chirurgien français Pierre Paul Broca, grâce à des autopsies poussées du cerveau, put montrer que les malades souffrant d'*aphasie* (l'incapacité de parler, ou de comprendre le langage) étaient atteints généralement de lésions dans la partie gauche du cerveau, une zone qu'on appelle maintenant la *circonvolution de Broca.*

Puis, en 1870, deux savants allemands, Gustav Fritsch et Eduard Hitzig, commencèrent à établir une liste des fonctions de supervision du cerveau en stimulant certaines parties à tour de rôle et en observant quels muscles répondaient. Un demi-siècle plus tard, cette technique fut améliorée par le physiologiste suisse Walter Rudolf Hess qui, pour ses travaux, partagea en 1949 le prix Nobel de médecine et de physiologie.

Ces méthodes ont permis de découvrir qu'une bande spécifique du cortex était impliquée dans la stimulation et la mise en mouvement des divers muscles volontaires. Cette bande a donc été baptisée *zone motrice.* Elle est, semble-t-il, caractérisée par une relation inverse par rapport au corps ; les parties supérieures de la zone motrice, vers le sommet du cerveau, stimulent les parties inférieures de la jambe ; au fur et à mesure que l'on descend dans la zone motrice, ce sont les muscles situés dans la partie supérieure de la jambe qui sont stimulés puis les muscles du torse, du bras et de la main, et finalement ceux du cou et de la main.

Derrière la zone motrice se trouve une autre section du cortex qui reçoit de nombreux types de sensations et qui, par conséquent, s'appelle la *zone sensorielle.* Comme pour la zone motrice, les parties de la zone sensorielle du cortex cérébral sont divisées en sections qui paraissent avoir une relation inverse par rapport au corps. Les sensations en provenance du pied se trouvent au sommet de la zone, suivies successivement par les sensations de la jambe, de la hanche, du tronc, du cou, du bras, de la main, des doigts, et tout en bas, de la langue. Les sections de la zone sensorielle affectées aux lèvres, à la langue et à la main sont (comme on pourrait s'en douter) plus importantes, proportionnellement à la taille de ces organes, que les sections affectées aux autres parties du corps.

Si l'on ajoute, à la zone motrice et à la zone sensorielle, les sections du cortex cérébral qui s'occupent en priorité de la réception des impressions en provenance des principaux organes sensoriels, l'œil et l'oreille, il reste encore une grande partie du cortex sans fonctions clairement établies.

C'est cette absence apparente d'attribution qui a donné naissance à l'idée, souvent entendue, que l'être humain « n'utilise qu'un cinquième de son cerveau ». Ceci, bien sûr, est inexact ; tout ce que l'on peut dire, c'est qu'un cinquième du cerveau humain possède des fonctions parfaitement définies. On pourrait tout aussi bien faire la comparaison avec une entreprise chargée de la construction d'un gratte-ciel, et dire qu'elle n'utilise qu'un cinquième de ses ouvriers parce que seuls ces derniers sont activement occupés à monter des barres d'acier, installer des gaines électriques, transporter des matériaux, etc. Cela reviendrait à faire bon compte des architectes, secrétaires, employés de bureaux d'études, contremaîtres, etc., qui sont également impliqués dans l'ouvrage. Par analogie, la majeure partie du cerveau est occupée à des tâches que l'on pourrait qualifier de « cols blancs » : collecter les données sensorielles, les analyser, décider quelles sont les informations superflues, quelles sont les actions nécessaires, et comment les mettre en œuvre. Le cortex cérébral possède des *zones associatives* distinctes – certaines affectées aux sensations auditives, d'autres aux sensations visuelles, etc.

Lorsqu'on tient compte de toutes ces zones associatives, il reste encore une région du cerveau sans fonction spécifique, et difficilement définissable. C'est la région située juste derrière le front, le *lobe préfrontal.* L'absence de toute fonction évidente lui a valu le surnom de *région silencieuse.* Des tumeurs ont parfois nécessité l'ablation d'une grande partie du lobe

préfrontal sans conséquences dramatiques pour le patient ; et pourtant, il ne s'agit sûrement pas là d'une masse inutile de tissu nerveux.

On pourrait même supposer qu'elle représente la portion la plus importante du cerveau si l'on considère qu'au cours du développement du système nerveux, on assiste à un entassement systématique et continuel de processus sophistiqués dans sa partie antérieure. Il se pourrait que le lobe préfrontal soit la région du cerveau la plus jeune et la plus véritablement « humaine ».

Dans les années trente, un chirurgien portugais, Antonio Egas Moniz, se dit qu'il serait peut-être possible d'aider, en dernière extrémité, un patient atteint d'une maladie mentale incurable en recourant à la solution drastique de l'ablation du lobe préfrontal. Peut-être le malade serait-il alors débarrassé d'une partie des associations qu'il avait faites durant son existence et qui apparemment l'affectaient, pour faire table rase et recommencer à vivre sur des bases plus saines avec le cerveau qui lui resterait. Cette opération, la *lobotomie préfrontale*, a été effectuée pour la première fois en 1935 ; dans un certain nombre de cas, on a pu noter une amélioration chez le sujet. Pour ses travaux, Moniz a partagé (avec W.R. Hess) le prix Nobel de médecine et de physiologie en 1949. Néanmoins, cette opération n'a jamais connu un grand succès et n'est presque plus pratiquée de nos jours. Trop souvent, le résultat est bien pire que la maladie.

Le cerveau est en réalité divisé en deux *hémisphères cérébraux* rattachés par un solide pont de substance blanche, le *corps calleux*. De fait, les hémisphères sont des organes séparés, unis dans l'action par les fibres nerveuses qui traversent le corps calleux et qui fonctionnent pour coordonner le tout. Mais les hémisphères demeurent néanmoins potentiellement indépendants.

Une situation analogue se présente dans le cas de nos yeux. Normalement, nos deux yeux fonctionnent comme une unité, mais si nous perdons un œil, l'autre peut continuer de remplir son rôle. De façon similaire, l'ablation d'un des hémisphères chez l'animal ne le prive pas entièrement de cerveau ; l'hémisphère restant apprend à prendre la relève de l'autre.

D'habitude, chaque hémisphère est en gros responsable d'un côté du corps : l'hémisphère cérébral gauche, du côté droit, l'hémisphère cérébral droit, du côté gauche. Si l'on ne touche pas aux deux hémisphères et que l'on sectionne le corps calleux, il y a perte de la coordination, et les deux moitiés du corps deviennent plus ou moins indépendantes. On crée ainsi une situation de cerveau dédoublé.

Quand on opère un singe de cette manière (en pratiquant une autre intervention sur le nerf optique pour s'assurer que les yeux sont bien connectés chacun à un seul hémisphère), on peut entraîner chaque œil à accomplir des tâches différentes. Un singe peut être dressé pour choisir une croix sur un cercle qui indique, par exemple, la présence de nourriture. Si on laisse l'œil gauche à découvert et qu'on recouvre l'œil droit pendant la période de dressage, seul l'œil gauche est utile dans l'expérience. Si on laisse l'œil droit à découvert et qu'on cache l'œil gauche, le singe n'a aucun souvenir de son dressage. Il cherche sa nourriture au petit bonheur. Si on dresse chaque œil à accomplir des tâches contradictoires, et si on découvre les deux yeux en même temps, le singe alterne les activités, puisque les hémisphères prennent poliment chacun leur tour.

Évidemment, dans une telle situation « à deux capitaines », il existe toujours un danger de conflit et de confusion. Afin d'éviter cela, un des hémisphères (presque toujours le gauche chez les êtres humains) est dominant, lorsque les deux sont normalement connectés. La circonvolution de Broca, qui contrôle le langage, se trouve dans l'hémisphère gauche, par exemple. La *zone gnostique*, une zone associative par excellence, une sorte de cour suprême, est également dans l'hémisphère gauche. Puisque l'hémisphère cérébral gauche guide l'activité motrice du côté droit du corps, il n'est pas surprenant qu'une majorité des gens soient droitiers (bien que même chez les gauchers, l'hémisphère gauche soit généralement aussi dominant). Dans les cas où la dominance entre la gauche et la droite n'est pas claire, le sujet est ambidextre, plutôt que clairement droitier ou gaucher ; il risque d'avoir des problèmes au niveau de la parole et de l'adresse manuelle.

Il est devenu de mode ces dernières années d'admettre que les deux moitiés du cerveau « pensent » différemment. L'hémisphère gauche, qui contrôle le langage, penserait logiquement, mathématiquement, pas à pas. L'hémisphère droit serait le siège de l'intuition, de la création artistique, de la pensée synthétique.

Mais le cerveau n'est pas uniquement cela. Il existe des zones de substance grise ensevelies sous le cortex cérébral. Ce sont les *ganglions de la base du cerveau* ; il y a également le *thalamus* (voir figure 17.2). Le thalamus fonctionne comme une sorte de centre d'accueil pour diverses sensations. Les plus violentes – comme la douleur, le froid et le chaud extrêmes, ou le toucher de choses rugueuses – sont filtrées. Les sensations plus douces en provenance des muscles – le toucher d'objets lisses, les températures moyennes – sont envoyées à la zone sensorielle du cortex cérébral. Tout se passe comme si les sensations douces pouvaient être confiées au cortex, qui peut les traiter judicieusement et envoyer une réponse après une période de réflexion plus ou moins prolongée. Les sensations fortes, cependant, qui doivent être traitées rapidement et pour lesquelles un temps de réflexion serait un luxe superflu, sont confiées plus ou moins automatiquement au thalamus.

Sous le thalamus se trouve l'*hypothalamus*, centre de nombreux processus qui contrôlent le corps. La régulation de l'appétit s'effectue là (voir chapitre 15), ainsi que celle de la température. De plus, c'est par l'intermédiaire de l'hypothalamus que le cerveau parvient à influer un tant soit peu sur la glande pituitaire (l'hypophyse ; voir également chapitre 15) ; voilà un bon exemple de la manière dont les centres nerveux du corps et les centres chimiques (les hormones) peuvent s'associer pour former une puissante organisation de contrôle.

En 1954, le physiologiste James Olds a découvert une autre fonction assez terrifiante de l'hypothalamus. Il existe dans ce dernier une zone qui, lorsqu'on la stimule, fait naître une intense sensation de plaisir. Une électrode placée dans le *centre de plaisir* d'un rat, reliée à un dispositif permettant à l'animal de déclencher lui-même la stimulation, peut être actionnée jusqu'à 8 000 fois par heure pendant plusieurs heures, ou même plusieurs jours, sans discontinuer ; l'animal en oublie la nourriture, le sexe et le sommeil. Il est évident que les choses agréables de la vie ne procurent du plaisir que dans la mesure où elles stimulent ce centre. Sa stimulation directe inhibe toutes les autres activités.

Il existe également dans l'hypothalamus une région qui intervient au niveau du cycle veille-sommeil ; une lésion dans cette région provoque chez l'animal un état de quasi-sommeil. Le mécanisme exact du fonctionnement de l'hypothalamus n'est pas très bien connu. Selon une théorie, il transmet des signaux au cortex, qui lui renvoie ses réponses, suivant un schéma de stimulation réciproque. L'état de veille continuel produit une rupture de la coordination entre les deux, les oscillations ne sont plus synchrones, et un état de sommeil s'installe chez l'individu. Un stimulus violent (un bruit fort, des secousses continues de l'épaule, ou bien encore l'interruption soudaine d'un bruit constant et régulier) réveille l'individu.

En l'absence de tels stimuli, la coordination entre l'hypothalamus et le cortex se rétablit au bout d'un laps de temps plus ou moins long, et l'état de sommeil cesse spontanément ; ou bien alors le sommeil devient si léger qu'un stimulus tout à fait ordinaire, si fréquent dans l'environnement, suffit à nous réveiller.

Pendant le sommeil surviennent les rêves – des données sensorielles plus ou moins dissociées du réel. Rêver est apparemment un phénomène universel ; ceux d'entre nous qui disent ne pas rêver ne se souviennent tout simplement pas de leurs rêves. Le physiologiste américain William Dement a étudié des sujets pendant leur sommeil en 1952, et a constaté des périodes de mouvements oculaires rapides qui duraient parfois plusieurs minutes (le *sommeil paradoxal*). Pendant cette période, la respiration, le rythme cardiaque et la tension montent à des niveaux comparables à ceux de l'état de veille. Ces périodes se produisent pendant un quart du temps de sommeil. Si l'on réveille quelqu'un au cours d'un tel épisode, il affirme généralement qu'il était en train de rêver. De plus, une personne que l'on dérange continuellement à ces moments-là commence à montrer des signes de fatigue psychologique ; les périodes de sommeil paradoxal se multiplient au cours des nuits suivantes comme pour compenser le manque de rêves.

Il semble donc que rêver soit une activité très importante dans le fonctionnement du cerveau. Des chercheurs suggèrent que rêver est un mécanisme qui permet au cerveau de récapituler les événements qui se sont produits dans la journée pour en filtrer ceux qui sont triviaux et répétitifs, et qui, sans cela, l'encombreraient et nuiraient ainsi à son efficacité. Le sommeil est la période naturelle pendant laquelle se produit cette activité, car le cerveau est alors déchargé des fonctions qui lui incombent en état de veille. S'il lui est impossible d'accomplir cette tâche pendant une période de sommeil (à cause d'interruptions trop fréquentes), il s'engorge tellement qu'il doit le faire pendant des périodes d'éveil, ce qui peut produire des hallucinations (autrement dit, des rêves éveillés), et d'autres symptômes déplaisants. On peut même se demander s'il ne s'agit pas là de la fonction essentielle du sommeil, puisqu'en fait le repos physique produit par le sommeil pourrait être facilement doublé par une période de détente pendant l'état de veille. Le sommeil paradoxal se produit même chez les nouveau-nés : ils y passent la moitié de leur temps, et pourtant, on voit mal à quoi ils pourraient bien rêver. Il se pourrait que le sommeil paradoxal facilite le développement du système nerveux. (On l'a également observé chez des mammifères autres que les êtres humains.)

LA MOELLE ÉPINIÈRE

Sous le cerveau se trouvent le cervelet, lui-même divisé en deux *hémisphères cérébelleux*, et la *tige cervicale*, qui se retrécit et rejoint la *moelle épinière* qui descend sur près de cinquante centimètres à l'intérieur de la colonne vertébrale.

La moelle épinière est faite de substance grise (au centre) et de substance blanche (à la périphérie) ; une série de nerfs s'y rattachent, reliés pour la plupart aux organes internes – le cœur, les poumons, le système digestif, etc.– des organes qui sont plus ou moins sous contrôle involontaire.

En général, si la moelle épinière est sectionnée, pour cause de maladie ou d'accident, la partie qui se trouve sous la lésion est pour ainsi dire déconnectée. Elle perd toute sensation et se trouve paralysée. Si la moelle est sectionnée dans la région du cou, la mort survient, car la poitrine est paralysée et, par là même, le fonctionnement des poumons. C'est pour cette raison qu'une rupture au niveau des cervicales est mortelle, et que la pendaison est une forme pratique d'exécution rapide. C'est la section de la moelle épinière, plus que la cassure de l'os, qui est mortelle.

Toute la structure du *système nerveux central*, c'est-à-dire le cérébrum, le cervelet, la tige cervicale, la moelle épinière, est parfaitement coordonnée. La substance blanche de la moelle épinière est faite de paquets de fibres nerveuses qui montent et descendent le long de la moelle et en assurent l'unité. Celles qui transportent les influx du cerveau vers le bas sont les *conduits descendants* et celles qui les transportent vers le cerveau sont les *conduits ascendants*

En 1964, des chercheurs du Metropolitan General Hospital de Cleveland ont réussi à isoler des cerveaux de singes rhésus et à les maintenir fonctionnels pendant des périodes allant jusqu'à dix-huit heures. Cette expérience a permis l'étude détaillée du métabolisme du cerveau, en comparant le contenu du milieu nutritif injecté dans les vaisseaux du cerveau isolé et celui du même milieu à sa sortie.

L'année suivante, cette équipe transplantait des têtes de chiens sur le cou d'autres chiens, en les connectant au système sanguin du receveur ; les têtes ainsi transplantées ont pu être maintenues en vie et fonctionnelles pendant plus de deux jours. Dès 1966, on réussissait à descendre la température de cerveaux de chiens à près de 0 °C pendant six heures, et à les faire revivre au point que se rétablisse nettement une activité chimique et électrique normale. Le cerveau est sans aucun doute un organe plus solide qu'on pourrait le croire à première vue.

Le mode d'action des nerfs

Ce ne sont pas seulement les différentes parties du système nerveux central qui sont raccordées par les nerfs, mais bien évidemment, le corps dans son ensemble ; celui-ci est ainsi placé tout entier sous le contrôle du système. Les nerfs envahissent les muscles, les glandes, la peau ; ils envahissent même la pulpe des dents (comme nous en faisons la douloureuse expérience quand nous sommes pris d'une rage de dents).

Les Anciens avaient déjà observé les nerfs, mais on s'est très longtemps mépris sur leur structure et leur fonctionnement. Récemment encore, on pensait qu'ils étaient creux et qu'ils servaient au transport d'un fluide. Des théories particulièrement compliquées, étayées par Galen, mentionnaient l'existence de trois différents fluides transportés respectivement par les veines, les artères et les nerfs. Le fluide des nerfs, généralement connu sous l'expression *esprits animaux*, est le plus raréfié des trois. La découverte de Galvani, selon laquelle les muscles et les nerfs peuvent être stimulés par une décharge électrique, a servi de fondement à toute une série d'études qui ont montré que le mode d'action des nerfs était lié à l'électricité – un fluide subtil, sans aucun doute, encore plus subtil que Galen ne pouvait l'imaginer.

Des travaux spécifiques sur le mode d'action des nerfs ont été engagés au début du XIXe siècle par le physiologiste allemand Johannes Peter Müller qui, entre autres, a pu démontrer que les nerfs sensoriels produisent toujours leurs sensations propres, quelle que soit la nature du stimulus. Ainsi, le nerf optique enregistre un éclair de lumière, qu'il soit stimulé par la lumière elle-même ou par la pression mécanique d'un coup de poing dans l'œil. (Dans ce dernier cas, « on voit des étoiles ».) Cet exemple montre à quel point nous n'avons en fait aucun contact direct avec la réalité du monde qui nous entoure, mais avec des stimuli bien spécialisés que le cerveau interprète d'habitude de manière correcte, mais peut également interpréter de façon erronée.

La connaissance des nerfs a fait un immense bond en avant en 1873, quand un physiologiste italien, Camillo Golgi, inventa un colorant cellulaire contenant des sels d'argent, très bien adapté à l'étude des nerfs, et permettant de les étudier dans leurs plus petits détails. Il put ainsi montrer que les nerfs sont composés de cellules séparées et distinctes, et que les processus qui interviennent dans une cellule, bien que très proches de ceux des cellules voisines, ne se confondent pas avec eux. Restait le minuscule passage de la synapse. Golgi, grâce à l'observation, confirma les hypothèses d'un anatomiste allemand, Wilhelm von Waldeyer, pour qui le système nerveux tout entier était formé de cellules nerveuses individuelles, ou neurones (hypothèses réunies sous l'expression *théorie neuronique*).

Golgi, néanmoins, ne soutenait pas lui-même la théorie neuronique. Cette tâche est revenue au neurologue espagnol Santiago Ramón y Cajal, qui, en 1889, en utilisant une version améliorée du colorant de Golgi, établit les connections entre les cellules de la substance grise et celles de la moelle épinière et prouva donc la théorie neuronique. Golgi et Ramón y Cajal, bien qu'en désaccord sur des points mineurs de leurs travaux, partagèrent le prix Nobel de médecine et de physiologie en 1906.

Les nerfs forment deux systèmes : le *sympathique* et le *parasympathique* (ces termes remontent aux notions quasi mystiques de Galen). Les deux systèmes ont une action sur presque tous les organes internes, et agissent par des effets opposés. Par exemple, les nerfs du système sympathique provoquent une accélération du rythme cardiaque, et ceux du parasympathique un ralentissement ; les nerfs du sympathique inhibent les sécrétions digestives, ceux du parasympathique les stimulent, etc. Ainsi, la moelle épinière, de concert avec les zones sous-cérébrales du cerveau, régulent automatiquement le fonctionnement des organes. Ce système automatique de contrôle a été étudié en détail par le physiologiste anglais

John Newport Langley dans les années 1890, et il l'a appelé le *système nerveux autonome.*

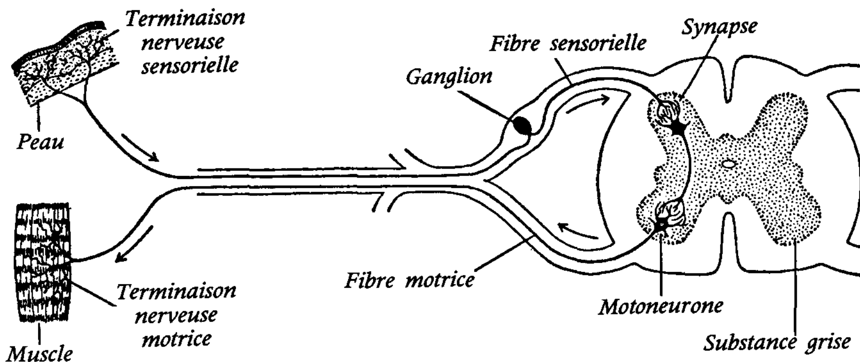
LES RÉFLEXES

Dans les années 1830, le physiologiste anglais Marshall Hall avait étudié un autre type de comportement qui semblait receler des aspects volontaires qui se sont en fait avérés tout à fait involontaires. Lorsque vous touchez accidentellement un objet brûlant avec votre main, celle-ci se retire instantanément. Si la sensation de chaleur devait remonter jusqu'au cerveau, être prise en considération et interprétée là-haut, pour finalement susciter le message approprié lui intimant l'ordre de se retirer, votre main serait bien brûlée avant que le message ne l'atteigne. La moelle épinière, malgré son absence totale d'intelligence, se charge automatiquement de tout l'épisode, et bien plus rapidement. C'est à Hall que l'on doit le nom de *réflexe* pour qualifier ce processus.

Le réflexe est produit par l'action coordonnée de deux nerfs ou plus, qui forment un *arc réflexe* (figure 17.3). L'arc réflexe le plus simple possible consiste en deux neurones, l'un sensoriel (qui envoie les sensations vers un *centre réflexe* dans le système nerveux central, d'habitude situé à un point quelconque de la moelle épinière), et un autre moteur (qui transporte les instructions de mouvement en provenance du système nerveux central). Les deux neurones peuvent être connectés par un ou plusieurs *neurones connecteurs.* Le neurologue anglais Charles Scott Sherrington a effectué des travaux sur ces arcs réflexes et sur leur fonctionnement dans le corps humain, qui lui ont valu de partager le prix Nobel de médecine et de physiologie en 1932. C'est Sherrington qui inventa le terme *synapse* en 1897.

Les réflexes provoquent une réponse si rapide et si certaine à un stimulus donné qu'ils offrent une méthode simple pour assurer l'intégrité et le bon fonctionnement du système nerveux. Le *réflexe rotulien* en est un exemple bien connu. Lorsqu'on croise les jambes, un coup brusque sous le genou provoque un mouvement rapide, une sorte de coup de pied – un phénomène dont la portée médicale a été pressentie pour la première fois

Figure 17.3. L'arc réflexe.



par le neurologue allemand Karl Friedrich Otto Westphall en 1875. Le réflexe rotulien n'est pas important en soi, mais c'est son absence qui peut être le signe d'un problème sérieux au niveau de la partie du système nerveux où se trouve cet arc réflexe.

Parfois, une lésion dans une partie du système nerveux central provoque l'apparition d'un réflexe anormal. Si l'on gratte la plante du pied, le réflexe normal fait se rapprocher les doigts de pied et les fait se plier vers le bas. Dans certains types de lésions du système nerveux central, le gros doigt de pied se plie vers le haut en réponse à ce stimulus, et les autres doigts s'écartent en se pliant vers le bas. Il s'agit du *réflexe de Babinski*, du nom du neurologue français qui l'a décrit en 1896.

Chez l'être humain, les réflexes sont parfois indubitablement subordonnés à la volonté. Ainsi, on peut accélérer son rythme respiratoire alors que le réflexe normal exigerait son ralentissement, ou vice versa. Dans les phylums inférieurs, les animaux sont beaucoup plus sous l'emprise de l'activité réflexe que l'homme, et chez eux, cette activité est beaucoup plus développée.

L'araignée qui tisse sa toile en offre un des meilleurs exemples. Ici, les réflexes produisent un ensemble de comportements si élaborés qu'on a du mal à n'y voir que la seule activité réflexe ; à la place, on parle habituellement de comportement *instinctif*. (A cause de l'utilisation abusive du mot *instinct* les biologistes préfèrent utiliser le terme *inné*.) L'araignée naît avec un système de câblage nerveux dont les interrupteurs ont été, pour ainsi dire, préréglés. Un stimulus donné la pousse à tisser sa toile, et chaque étape du processus devient à son tour un stimulus qui détermine la réponse suivante.

Lorsqu'on regarde la toile incroyablement complexe de l'araignée, tissée avec une précision étonnante et une efficacité tout entière destinée au but recherché, il est quasi impossible de croire qu'elle a été faite sans une volonté intelligente pour la guider. Et pourtant, le fait même que cette tâche si complexe s'accomplisse aussi parfaitement et toujours de manière identique offre la preuve patente que l'intelligence n'a rien à voir là-dedans. L'intelligence consciente, avec ses hésitations et ses incertitudes liées aux alternatives et inhérentes à l'action délibérée, s'accompagne inévitablement d'imperfections et de variations d'une construction à l'autre.

Avec le développement de l'intelligence, les animaux ont de plus en plus tendance à perdre leurs instincts et leurs dispositions innées. Ils perdent sans doute, par le fait même, certaines qualités inestimables. Une araignée peut tisser sa toile parfaitement la première fois, bien qu'elle n'ait jamais vu ni une autre araignée, ni même une toile. L'être humain, quant à lui, naît dépourvu de toute compétence particulière, et sans ressources. Un nouveau-né sait automatiquement téter, pleurer s'il a faim, mais pas grand-chose de plus. Tous les parents savent combien de patience et de sollicitude il leur faut déployer avant qu'un enfant n'apprenne, après maints efforts, les gestes les plus simples de la vie. Une araignée ou n'importe quel autre insecte, bien que possédant la perfection innée, ne peut en dévier. L'araignée tisse une toile splendide, mais s'il s'avère que cette toile prédéterminée remplisse mal sa fonction, elle est incapable d'apprendre à tisser une autre forme de toile. Un enfant, en revanche, tire grand profit de l'absence de cette perfection innée. Il apprendra lentement et n'atteindra sans doute jamais la perfection, mais il pourra lui-même choisir les domaines dans lesquels il ne sera qu'imparfait. Ce

que l'être humain perd en confort et en sécurité, il le gagne sous la forme d'une flexibilité quasi illimitée.

Cependant, des travaux récents soulignent le fait qu'il n'existe pas toujours une frontière nette entre les comportements inné et appris, non seulement dans le cas du feedback humain, mais aussi chez les animaux inférieurs. A première vue, il semble par exemple que les poussins ou les canetons suivent leur mère d'instinct au sortir de la coquille. Une observation plus attentive démontre le contraire.

Leur instinct dicte en fait de suivre, non pas leur mère, mais simplement tout objet caractérisé par une certaine forme ou une certaine couleur, et doué de mouvement. N'importe quel objet produisant cette sensation pendant la période néo-natale est suivi par le petit qui l'adopte dès lors comme sa mère. Il peut bien entendu s'agir réellement de la mère ; de fait, c'est pratiquement toujours le cas, mais ce n'est pas obligatoire ! Autrement dit, suivre est instinctif, mais la « mère » qui est suivie est « apprise ». (Une grande partie du mérite de cette découverte revient au remarquable naturaliste autrichien Konrad Zacharias Lorenz. Ce dernier, au cours de travaux datant maintenant d'il y a près de trente ans, était suivi régulièrement dans tous ses déplacements par un troupeau d'oisillons piaillant à qui mieux mieux.)

L'établissement des éléments fixes d'un comportement en réponse à un stimulus donné rencontré à une certaine période de la vie s'appelle l'*imprégnation*. La période exacte à laquelle ce phénomène se produit s'appelle la *période critique*. Chez les poussins, la période critique de l'*imprégnation maternelle* se situe entre treize et seize heures après l'éclosion. Chez le chiot, cette période est comprise entre trois et sept semaines, pendant lesquelles les stimulations qu'il est à même de ressentir impriment les diverses conduites caractéristiques du comportement canin normal.

L'imprégnation constitue la forme la plus primitive d'apprentissage, qui est tellement automatique, se produit en si peu de temps, et dans des conditions si générales, qu'on la confond facilement avec l'instinct.

Une raison logique de l'imprégnation est qu'elle gratifie l'individu d'un degré nécessaire de flexibilité. Si le poussin naissait doté d'une faculté instinctive lui permettant de reconnaître sa vraie mère et l'obligeant à ne suivre qu'elle, et si pour une raison quelconque sa vraie mère était absente pendant le premier jour de sa vie, la petite créature serait totalement démunie. Dans l'état actuel des choses, la question maternelle reste en suspens pendant seulement quelques heures, et le petit poussin peut s'imprégner de n'importe quelle poule aux alentours et choisir ainsi une mère adoptive.

L'INFLUX ÉLECTRIQUE

Comme je l'ai expliqué plus haut, ce sont les expériences de Galvani à la fin du XVIIIe siècle qui ont, les premières, établi un rapport entre l'électricité et le fonctionnement des muscles et des nerfs.

Les propriétés électriques des muscles ont suscité une application médicale étonnante, grâce aux travaux du physiologiste néerlandais Willem Einthoven. En 1903, il a mis au point un galvanomètre très sensible, assez sensible pour mesurer les minuscules fluctuations de potentiel du cœur quand il bat. Dès 1906, Einthoven était capable

d'enregistrer les pics et les creux de ce potentiel (l'enregistrement s'appelle un *électrocardiogramme*) et de les corréler avec différents types de maladies cardiaques.

On pensait à l'époque que les propriétés électriques de l'influx nerveux étaient provoquées par, et se propageaient grâce à des modifications chimiques dans le nerf. Au XIXe siècle, les travaux du physiologiste allemand Emil Du Bois-Reymond ont apporté la démonstration expérimentale de l'exactitude de cette hypothèse, par la détection d'infimes courants électriques dans des nerfs stimulés.

Grâce aux instruments électroniques modernes, la recherche sur les propriétés électriques des nerfs a fait d'immenses progrès. En plaçant de minuscules électrodes à différents endroits d'une fibre nerveuse, et en détectant les différences de potentiel électrique au moyen d'un oscilloscope, il est possible de mesurer l'intensité, la durée, la vitesse de propagation, etc., d'un influx nerveux. Les travaux des physiologistes américains Joseph Erlanger et Herbert Spencer Gasser sur ce sujet leur ont valu en 1944 le prix Nobel de médecine et de physiologie.

Si l'on applique à une seule cellule nerveuse de petites impulsions électriques d'intensité croissante, on n'enregistre aucune réponse jusqu'à un certain point. Puis, tout à coup, la cellule réagit : une impulsion est créée, qui se propage le long de la fibre. La cellule a un seuil ; on ne note aucune réaction de la cellule en dessous de ce seuil, et tout stimulus au-dessus de ce seuil provoque seulement une réponse caractérisée par une certaine intensité. En d'autres termes, la réponse est du mode « tout ou rien ». Et la nature de l'impulsion provoquée par le stimulus semble être identique pour tous les nerfs.

Comment un processus aussi simple, de type oui-non, peut-il déclencher des sensations aussi complexes que la vue, par exemple, ou provoquer des réponses aussi compliquées que celle des doigts d'un violoniste ? Il semblerait que dans un nerf, le nerf optique par exemple, constitué d'un très grand nombre de fibres individuelles, chacune de celles-ci se *décharge* à un rythme qui lui est propre, à des intervalles lents ou en succession rapide, créant ainsi des combinaisons extrêmement complexes et changeantes en fonction des modifications globales du stimulus. (Pour ses travaux dans ce domaine, le physiologiste anglais Edgar Douglas Adrian a partagé, avec Sherrington, le prix Nobel de médecine et de physiologie en 1932.) Ces combinaisons changeantes sont continuellement *déchiffrées* par le cerveau et interprétées. Mais nous ne savons pratiquement rien sur la nature des processus d'interprétation, et nous ignorons encore comment ces combinaisons sont traduites en réponses telles que la contraction musculaire ou la sécrétion glandulaire.

La dépolarisation de la cellule nerveuse en soi dépend apparemment du transport des ions à travers la membrane cellulaire. Normalement, l'intérieur de la cellule contient un excédent relatif d'ions potassium, alors que l'extérieur est caractérisé par un excédent d'ions sodium. D'une façon ou d'une autre, la cellule maintient les ions potassium à l'intérieur et les ions sodium à l'extérieur pour éviter que leurs concentrations respectives des deux côtés de la cellule ne s'équilibrent. On pense aujourd'hui qu'une sorte de *pompe à sodium* à l'intérieur de la cellule refoule les ions sodium au fur et à mesure qu'ils arrivent. En tout cas, il existe une différence de potentiel d'environ un dixième de volt le long de la membrane cellulaire, l'intérieur étant de charge négative par rapport à l'extérieur.

Lorsque la cellule nerveuse est stimulée, la différence de potentiel à la surface de la membrane s'effondre, provoquant la dépolarisation de la cellule. Il faut deux millièmes de seconde pour que la différence de potentiel se rétablisse ; et pendant cet intervalle, le nerf ne réagit à aucun autre stimulus. Il s'agit là de la *période réfractaire*.

Lorsque la cellule se dépolarise, l'influx nerveux voyage le long de la fibre en une série de décharges, chaque section successive de la fibre excitant à son tour la section voisine. L'influx ne peut se propager que dans un sens, car la section qui vient de se dépolariser doit attendre un certain temps avant de se repolariser à nouveau.

Des travaux montrant, comme je viens de le décrire, une corrélation entre le fonctionnement du nerf et la perméabilité des ions ont valu le prix Nobel 1963 à deux physiologistes anglais, Alan Lloyd Hodgkin et Andrew Fielding Huxley, et à un physiologiste australien, John Carew Eccles.

Mais que se passe-t-il lorsque l'influx qui se propage le long de la fibre nerveuse atteint une synapse – étroit passage entre une cellule nerveuse et la suivante ? Apparemment, l'influx nerveux implique également la production d'une substance chimique qui peut franchir ce passage et provoquer la dépolarisation de la cellule suivante. De cette manière, l'influx peut se propager d'une cellule à l'autre.

L'une des substances chimiques certainement impliquées dans le fonctionnement des nerfs est l'adrénaline. Elle agit sur les nerfs du système sympathique, qui inhibe l'activité du système digestif et accélère les rythmes respiratoire et cardiaque. Lorsque la colère ou la peur excitent les glandes adrénalines, provoquant ainsi la sécrétion de l'hormone, l'action de cette dernière sur les nerfs sympathiques produit une augmentation soudaine de la circulation sanguine dans tout le corps, transportant une quantité accrue d'oxygène vers les tissus ; et puisque l'hormone ralentit la digestion par la même occasion, il y a économie d'énergie pendant l'état d'urgence.

Les psychologues et officiers de police américains John Augustus Larsen et Leonard Keeler ont mis à profit ces découvertes en 1921 et inventé un appareil permettant la détection des modifications de la tension artérielle, du pouls, du rythme respiratoire et de la transpiration provoquées par l'émotion. Cet appareil, le *polygraphe*, détecte le stress causé par un mensonge, qui recèle toujours chez n'importe quel individu normal une part de crainte d'être pris au piège, et fait donc entrer l'adrénaline en jeu. Bien que loin d'être infaillible, le polygraphe connaît une grande notoriété en tant que *détecteur de mensonges*.

Dans le cours normal des choses, les terminaisons nerveuses du système sympathique sécrètent un produit tout à fait similaire à l'adrénaline, la *noradrénaline*. Cette substance sert à transporter l'influx nerveux à travers les synapses, en transmettant les messages par stimulation des terminaisons nerveuses situées de l'autre côté du passage.

Au tout début des années vingt, le physiologiste anglais Henry Dale et le physiologiste allemand Otto Loewi (qui devaient se partager le prix Nobel de physiologie et de médecine en 1930) ont étudié une substance qui remplissait la même fonction pour la plupart des nerfs autres que ceux du système sympathique. Cette substance s'appelle l'*acétylcholine*. On pense aujourd'hui qu'elle intervient non seulement au niveau des synapses, mais aussi au niveau de la conduction de l'influx nerveux le long de la fibre

elle-même. Peut-être l'acétylcholine joue-t-elle aussi un rôle au niveau de la pompe à sodium. En tout cas, cette substance paraît se former momentanément dans la fibre nerveuse pour ensuite être dégradée rapidement par une enzyme, la *cholinestérase*. Toute inhibition de l'action de la cholinestérase provoque une rupture du cycle et bloque la transmission de l'influx nerveux. Les substances mortelles qui constituent l'arsenal de la *guerre chimique* sont des inhibiteurs de la cholinestérase. En interrompant la conduction des influx nerveux, elles peuvent arrêter le cœur et provoquer la mort en quelques secondes. Leur utilisation guerrière est évidente, mais elles peuvent servir dans un but moins immoral comme insecticides.

Les anesthésiques locaux interfèrent de façon moins drastique avec la cholinestérase, et interrompent (momentanément) les influx nerveux associés à la douleur.

Grâce aux courants électriques impliqués dans l'influx nerveux, il est possible de « lire » l'activité cérébrale, d'une certaine manière, bien que personne à ce jour n'ait été capable d'interpréter ce que les ondes du cerveau pouvaient bien signifier. En 1929, un psychiatre allemand, Hans Berger, publia les résultats de travaux antérieurs dans lesquels il appliquait des électrodes en divers endroits du crâne, ce qui lui avait permis de détecter les ondes rythmiques d'une certaine activité électrique.

Berger baptisa le rythme le plus prononcé *onde alpha*. Dans l'onde alpha, le potentiel varie d'environ vingt microvolts à une fréquence avoisinant dix cycles par seconde. L'onde alpha est particulièrement forte lorsque le sujet est au repos, les yeux fermés. Lorsqu'il ouvre les yeux et regarde une source lumineuse, à l'exclusion de tout autre stimulus, l'onde alpha persiste. Mais s'il regarde un environnement habituel, et donc diversifié, l'onde alpha disparaît, ou plutôt est noyée par d'autres rythmes plus accentués. Après un certain temps, si aucune nouvelle stimulation visuelle ne se produit, l'onde alpha réapparaît. Les autres types d'ondes portent bien sûr des noms tels que *bêta*, *delta* et *thêta*.

Les *électro-encéphalogrammes* (« messages électriques du cerveau » ou, en abrégé, EEG) ont été depuis l'objet d'études très poussées, et montrent que chaque individu possède son propre tracé, qui varie selon qu'il est excité ou qu'il dort. Bien que l'électro-encéphalogramme ne permette pas encore de « lire les pensées », ou de suivre les mécanismes de l'intellect, il est utile dans le diagnostic des principales affections mentales, particulièrement de l'épilepsie. Il permet aussi parfois de localiser des lésions ou des tumeurs cérébrales.

Dans les années soixante, des ordinateurs spécialement programmés ont été appelés à la rescousse. L'hypothèse de départ était la suivante : si l'on soumet un sujet à une petite modification de son environnement, celle-ci provoquera une réponse dans son cerveau qui se traduira par une petite altération de son EEG au moment où la modification aura lieu. Mais le cerveau sera occupé par bon nombre d'autres activités, et la petite modification de l'EEG ne sera pas perceptible. Néanmoins, si l'on répète la procédure plusieurs fois, on peut programmer un ordinateur pour calculer la moyenne des tracés et trouver ainsi une différence systématique.

Dès 1964, le psychologue américain Manfred Clynes publiait les résultats d'analyses précises permettant d'affirmer, grâce au seul tracé de l'EEG, quelle couleur un sujet regardait. Le neurophysiologiste anglais William Grey Walter démontrait également l'existence d'un signal cérébral

apparemment caractéristique des processus de l'apprentissage. Il apparaît lorsque le sujet étudié a tout lieu de penser qu'on va lui présenter un stimulus qui nécessitera de sa part une réflexion ou une action. Walter l'appelle *l'onde d'attente* et souligne qu'elle n'existe pas chez l'enfant de moins de trois ans et chez certains psychotiques. Le phénomène inverse, qui consiste à provoquer des réactions spécifiques par des stimulations électriques directes du cerveau, a également été signalé en 1965. José Manuel Rodriguez Delgado de l'université de Yale a fait marcher, grimper, bâiller, s'accoupler, des animaux de son laboratoire et changer leur activité psychique sur simple ordre, au moyen de stimulations électriques transmises par signal radio. Son expérience la plus spectaculaire : arrêter un taureau en plein galop et le faire s'en retourner en trottant gentiment.

Le comportement humain

A la différence des phénomènes physiques, comme le cours des planètes ou les propriétés de la lumière, on n'a jamais pu enfermer le comportement des êtres vivants dans le cadre rigoureux des lois naturelles, et on ne le pourra peut-être jamais. Beaucoup prétendent que l'étude du comportement humain ne deviendra jamais une vraie science, dans la mesure où l'on ne pourra jamais expliquer ou prédire un comportement dans une situation donnée sur la base de lois naturelles et universelles. Et pourtant la vie ne fait pas exception au cadre des lois de la nature, et l'on peut soutenir que n'importe quel comportement d'un être vivant pourrait s'expliquer dans son ensemble si l'on en connaissait tous les facteurs. Mais c'est justement là que le bât blesse. Il est peu probable qu'on connaîtra jamais tous ces facteurs ; ils sont trop nombreux et trop complexes. Il ne faut cependant pas désespérer de ne jamais pouvoir faire progresser notre savoir sur nous-mêmes. Il nous reste encore beaucoup à faire pour mieux comprendre notre esprit dans toute sa complexité, et quand bien même nous n'arriverions jamais au bout du chemin, il faudrait toujours espérer qu'il est encore long et fructueux.

Le sujet est non seulement particulièrement compliqué, mais son étude systématique n'a commencé qu'il y a peu de temps. La physique est devenue majeure en 1600, et la chimie en 1775, mais le domaine beaucoup plus complexe de la *psychologie expérimentale* ne remonte qu'à 1879, lorsque le physiologiste allemand Wilhelm Wundt a mis sur pied le premier laboratoire entièrement destiné à l'étude du comportement humain. Wundt s'intéressait principalement aux sensations et à la manière dont les humains perçoivent les détails de l'univers qui les entoure.

A peu près au même moment, l'étude d'un domaine bien défini du comportement humain – celui de l'être humain en tant qu'engrenage dans l'entreprise industrielle – a vu le jour. En 1881, l'ingénieur américain Frederick Winslow Taylor s'est lancé dans une étude sur la mesure du temps requis pour l'accomplissement de certaines tâches, et a développé des méthodes d'organisation du travail permettant de raccourcir ce temps. Il a été le premier *expert en rendement*, et par voie de conséquence, s'est fait détester des ouvriers.

Mais plus on avance dans l'étude du comportement humain, soit dans les conditions bien contrôlées d'un laboratoire, soit de manière empirique dans le cadre d'une usine, plus on a l'impression de s'attaquer au mécanisme délicat d'une montre au moyen d'une clef à molette.

Chez les organismes simples, on peut observer des réponses directes, automatiques, appelées *tropismes* (du grec « tourner »). Les plantes sont le siège de *phototropismes* (réactions à la lumière), d'*hydrotropismes* (réactions à l'eau) et de *chimiotropismes* (réactions à diverses substances chimiques). Le chimiotropisme est aussi caractéristique de nombreux animaux, des protozoaires aux fourmis. Certains lépidoptères peuvent être attirés par un effluve à plus de trois kilomètres. Le fait que les tropismes soient complètement automatiques est démontré par le lépidoptère qui, victime de son phototropisme, ira jusqu'à voler dans la flamme d'une bougie.

Les réflexes dont nous avons parlé précédemment dans ce chapitre n'apportent guère de progrès par rapport aux tropismes, et l'imprégnation, également abordée, représente une forme d'apprentissage, mais tellement mécanique qu'elle mérite à peine cette définition. Et pourtant, ni les réflexes ni l'imprégnation ne peuvent être considérés comme relevant du seul domaine des animaux inférieurs ; les êtres humains en ont reçu leur lot.

LES RÉPONSES CONDITIONNÉES

Dès la naissance, le nouveau-né attrape le doigt qu'on applique contre la paume de sa main et le serre fermement, il tète le sein qu'on présente à ses lèvres. L'importance de tels instincts, qui l'empêchent de tomber et de mourir de faim, est évidente.

Il semble presque inévitable que le nouveau-né soit aussi sujet à l'imprégnation. La question ne peut pas être étudiée bien sûr par l'expérimentation, mais elle peut être approfondie grâce à des observations fortuites. Les enfants qui, à l'âge où ils commencent à babiller, ne sont pas exposés aux sons de la parole humaine, sont pratiquement incapables de parler plus tard, ou alors dans des limites très restreintes. Les enfants qui, élevés dans des institutions où ils sont nourris de façon satisfaisante et sont bien entretenus physiquement, mais ne reçoivent ni chaleur ni affection, donnent des résultats généralement catastrophiques. Leur développement physique et mental est fort retardé, et beaucoup d'entre eux meurent pour la simple raison qu'il leur manque une *présence maternelle* – terme qui recouvre peut-être le manque de stimuli adéquats causant l'imprégnation de certaines données nécessaires à l'acquisition de comportements spécifiques. De même, les enfants indûment privés des stimuli qui accompagnent la présence d'autres enfants pendant certaines périodes critiques de l'enfance développent des personnalités qui présentent d'une manière ou d'une autre des troubles sérieux.

Bien sûr, on peut toujours soutenir que les réflexes et l'imprégnation ne concernent que l'enfance. Arrivé à l'âge adulte, l'homme est un être rationnel répondant d'une manière qui n'est plus mécanique. Mais est-ce bien le cas ? Autrement dit, sommes-nous dotés d'un *libre arbitre* (comme il nous plaît de le croire) ? Ou bien notre comportement est-il absolument déterminé, d'une manière ou d'une autre, par les stimuli, comme le taureau dans l'expérience de Delgado que j'ai décrite plus haut ?

On peut soutenir l'existence du libre arbitre sur les plans philosophique ou théologique, mais je ne connais personne qui ait jamais pu trouver le moyen de le prouver expérimentalement. Démontrer le *déterminisme*, c'est-à-dire le contraire du libre arbitre, n'est pas chose aisée non plus. Certains, pourtant, s'y sont essayés, dont le plus connu se trouve être le physiologiste russe Ivan Petrovitch Pavlov.

Au départ, Pavlov était surtout intéressé par les mécanismes de la digestion. Il put montrer, dans les années 1880, que les liquides gastriques étaient sécrétés dans l'estomac dès qu'on plaçait de la nourriture sur la langue d'un chien ; l'estomac sécrétait ces liquides même si la nourriture ne l'atteignait jamais. Mais si l'on sectionnait le *nerf pneumogastrique* (qui relie le bulbe rachidien à différentes parties du canal alimentaire) près de l'estomac, les sécrétions cessaient. Les travaux qu'il a accomplis sur la physiologie de la digestion ont valu à Pavlov le prix Nobel de physiologie et de médecine en 1904. Mais comme bien d'autres lauréats du prix Nobel (particulièrement Ehrlich et Einstein), Pavlov a fait d'autres découvertes qui ont fait oublier celle pour laquelle il a effectivement reçu le prix.

Il décida d'étudier la nature automatique, ou réflexe, des sécrétions, en prenant la salive comme exemple commode et facilement observable. La vue ou l'odeur de la nourriture fait saliver le chien (les humains aussi d'ailleurs). Pavlov a décidé de faire sonner une cloche chaque fois qu'il plaçait de la nourriture devant un chien. Après un certain temps, entre vingt et quarante associations de la sorte, le chien salivait lorsqu'il entendait la cloche même en l'absence de nourriture. Une association venait d'être créée. L'influx nerveux qui transportait le son de la cloche vers le cerveau était devenu l'équivalent de celui représentant la vue ou l'odeur de la nourriture.

En 1903, Pavlov inventa le terme *réflexe conditionné* pour qualifier ce phénomène ; la salivation était une *réponse conditionnée*. Qu'il le veuille ou non, le chien salivait au son de la cloche comme il l'aurait fait en voyant la nourriture. Bien sûr, la réponse conditionnée pouvait être effacée – par exemple, en lui retirant la nourriture lorsque sonnait la cloche et, à la place, en lui envoyant une faible décharge électrique. Après un certain temps, le chien ne salivait plus mais tressaillait au son de la cloche, même quand on ne lui administrait plus de décharge électrique.

Plus tard, Pavlov est parvenu à créer certains choix chez des chiens en associant leur nourriture à une image lumineuse circulaire, et une décharge électrique à une image lumineuse en forme d'ellipse. Les chiens étaient parfaitement capables de faire la distinction, mais en variant l'image elliptique jusqu'à la rendre presque circulaire, cette distinction devenait quasi impossible. Finalement, le chien ne parvenait plus à décider, et montrait des signes de ce que l'on pourrait appeler une *dépression nerveuse*.

Les expériences de conditionnement sont donc devenues de puissants outils en psychologie. Grâce à elles, les animaux parlent presque à l'expérimentateur. Cette technique a rendu possibles des études approfondies sur certains animaux, afin de mieux connaître leur vitesse d'apprentissage, leurs instincts, leur acuité visuelle, leur capacité à distinguer les couleurs, etc. De toutes ces recherches, celles du naturaliste autrichien Karl von Frisch ressortent nettement. Von Frisch a dressé des abeilles de telle manière qu'elles volent vers certains plats situés en différents endroits pour trouver leur nourriture, et il s'est rendu compte

que ces petites bêtes retournaient peu après dans leur ruche pour renseigner précisément leurs petits camarades sur ces lieux. Grâce à ces expériences, Frisch à découvert que les abeilles savent distinguer certaines couleurs – dont l'utraviolet, mais pas le rouge – et qu'elles se les communiquent entre elles en dansant sur les rayons de miel ; que la nature et la vigueur de la danse donnent la direction et la distance de la nourriture par rapport à la ruche, et même des renseignements sur la quantité de l'approvisionnement ; que les abeilles, enfin, parviennent à reconnaître la direction grâce à la polarisation de la lumière dans le ciel. Les découvertes fascinantes de von Frisch sur le langage des abeilles ont ouvert tout un nouveau champ d'études sur le comportement animal.

En théorie, tout apprentissage renferme une certaine composante de réponses conditionnées. Lorsqu'on apprend à taper à la machine, par exemple, on commence par regarder le clavier, puis on substitue progressivement certains mouvements automatiques des doigts à la sélection visuelle de la touche demandée. Ainsi, penser à la lettre *k* s'accompagne d'un mouvement particulier du majeur de la main droite ; penser à *les* commande à l'annulaire droit, au majeur gauche puis à l'annulaire gauche de frapper successivement sur les touches correspondantes du clavier. Ces réponses ne font aucunement appel au conscient. Après un certain entraînement, la dactylo doit même s'arrêter pour se rappeler où sont les lettres. Personnellement, je tape très bien à la machine et d'une manière totalement mécanique, et si quelqu'un me demandait où se trouve le *f* par exemple, la seule façon que j'aurais de répondre (à part bien sûr en regardant le clavier) serait de bouger mes doigts comme si je tapais et d'essayer d'en prendre un sur le fait au moment où il tape un *f.* Seuls mes doigts connaissent le clavier ; pas mon esprit conscient.

Le même principe s'applique à des processus d'apprentissage plus complexes, comme lire ou jouer du violon. Pourquoi, après tout, le mot CRAIE écrit en noir sur cette page évoque-t-il automatiquement l'image d'un bâton blanc et faisant un bruit caractéristique sur un tableau noir, le tout représentant un mot ? Inutile d'épeler chaque lettre ou de faire de grands efforts de mémoire, ou de faire appel au raisonnement pour expliciter le message contenu dans le mot ; après une certaine période de conditionnement, le symbole est automatiquement associé à l'objet.

Dans les premières décades de ce siècle, le psychologue américain John Broadus Watson a échafaudé toute une théorie du comportement humain, qui s'appelle le *behaviorisme,* sur la base du conditionnement. Watson est même allé jusqu'à suggérer que l'homme n'a aucun contrôle conscient sur la façon dont il agit ; c'est le conditionnement qui détermine tout. Bien que sa théorie ait connu un certain succès à une époque, elle n'a jamais fait beaucoup d'émules parmi les psychologues. Tout d'abord, même si la théorie est fondamentalement correcte – si le comportement est dicté seulement par le conditionnement – le behaviorisme n'apporte pas d'éclaircissements sur les aspects du comportement humain qui sont d'un grand intérêt pour nous, tels que la créativité, les dons artistiques, le sens du bien et du mal. Il serait impossible d'identifier toutes les influences qui nous conditionnent, et de les relier à notre pensée et à nos croyances de manière normative ; or, quelque chose que l'on ne peut pas mesurer se prête difficilement à l'analyse scientifique.

Deuxièmement, quelle peut bien être la part du conditionnement dans un processus tel que l'intuition ? L'esprit met tout à coup côte à côte deux

idées ou deux événements qui n'avaient jusque-là jamais eu aucun lien, apparemment par hasard, et crée une idée ou une réponse entièrement nouvelle.

Les chiens et les chats, lorsqu'ils doivent résoudre certains problèmes (par exemple, trouver comment fonctionne le levier qui commande l'ouverture d'une porte) font appel à la technique des essais et des erreurs. L'animal fait des allées et venues frénétiques, sans rime ni raison, jusqu'à ce qu'il déclenche accidentellement l'ouverture de la porte. Si on le fait recommencer, un vague souvenir de la procédure correcte le fera peut-être ouvrir la porte un peu plus rapidement, et puis encore plus rapidement à l'essai suivant, jusqu'à ce qu'il actionne immédiatement le levier. Plus l'animal est intelligent, moins il lui faut de tentatives pour passer de la technique des essais et des erreurs à l'action dirigée et efficace.

Dès que nous abordons l'homme, la mémoire n'est plus défaillante. Vous aurez peut-être tendance à chercher une pièce de monnaie en jetant des coups d'œil un peu au hasard sur le plancher, mais une expérience passée vous guidera sans doute à la rechercher dans un endroit où vous l'avez déjà trouvée, ou à regarder à l'endroit d'où provenait le son lorsqu'elle est tombée, ou à entamer une recherche systématique sur le plancher. De la même manière, si vous étiez enfermé, vous essayeriez peut-être de vous échapper en touchant les murs au hasard ; mais vous sauriez à quoi ressemble une porte et vous concentreriez vos efforts sur elle.

En bref, l'homme réduit la technique des essais et des erreurs au strict minimum en faisant appel à ses années d'expérience, et en transformant ses idées en actions. Au cours de la recherche d'une solution, l'homme peut ne rien faire d'autre qu'agir en pensée. C'est cette méthode d'essais et d'erreurs au-dessus de toute contingence matérielle que nous appelons la *raison* et cette dernière n'est pas entièrement l'apanage de l'espèce humaine.

Les singes, dont les comportements sont plus simples et plus mécaniques que les nôtres, font preuve de jugement spontané, proche de ce que nous appelons la raison. Le psychologue allemand Wolfgang Köhler, pris au dépourvu par la déclaration de guerre en 1914 dans une des colonies allemandes d'Afrique, trouva des illustrations frappantes de ce phénomène au cours d'expériences célèbres avec des chimpanzés. Dans l'une d'elles, un jeune chimpanzé, qui avait essayé en vain d'atteindre des bananes à l'aide d'un bâton trop court, ramassa soudain un autre morceau de bambou que l'expérimentateur avait laissé traîner à proximité, mit les deux morceaux bout à bout, et put ainsi atteindre le fruit convoité. Une autre fois, un chimpanzé empila deux boîtes l'une sur l'autre pour cueillir des bananes accrochées à une branche au-dessus de sa tête. Ces activités n'avaient été précédées d'aucune séance de dressage ou d'expériences qui auraient pu former des associations chez l'animal ; apparemment, il s'agissait là de purs éclairs d'inspiration.

Köhler pensait que l'apprentissage impliquait l'ensemble d'un processus, plutôt que des éléments épars. Il est l'un des fondateurs d'une école de pensée, la *Gestalt.*

Les chimpanzés et les autres grands singes ont une apparence et un comportement si humains que de nombreuses expériences ont été tentées dans lesquelles de jeunes singes sont élevés avec des enfants dans le but de comparer leurs progrès respectifs. Au début, grandissant plus rapidement, les jeunes singes dépassent leurs alter ego humains. Mais dès

que les enfants apprennent à parler, les singes sont définitivement dépassés. Il leur manque l'équivalent de la circonvolution de Broca.

Cependant, dans la nature, les chimpanzés communiquent non seulement grâce à un petit nombre de sons, mais également par le geste. Beatrice et Allen Gardner, de l'université du Nevada, ont eu l'idée en 1966 d'apprendre à une femelle chimpanzé d'un an et demi, appelée Washoe, un langage basé sur des signes. Les résultats dépassèrent leurs espérances. Washoe apprit plusieurs dizaines de symboles, les utilisant correctement, et les comprenant facilement.

D'autres chercheurs renouvelèrent ces expériences avec d'autres chimpanzés – avec de jeunes gorilles également. Et la controverse s'installa. Est-ce que les singes communiquaient de façon créative, ou répondaient-ils mécaniquement par réflexes conditionnés ?

Ceux qui avaient prodigué leur enseignement aux singes ne comptaient plus les anecdotes sur leurs petits protégés, racontant que ces derniers inventaient souvent de nouvelles combinaisons de symboles, mais ces histoires sont rejetées par les critiques pour leur manque de preuves ou d'objectivité. Il ne fait aucun doute que la controverse continuera bon train.

Il est apparu que la puissance du conditionnement est bien plus grande qu'on ne l'avait supposé, même chez l'être humain. On pensait depuis longtemps que certaines fonctions du corps – telles que le pouls, la tension artérielle, les contractions intestinales – étaient essentiellement sous le contrôle du système nerveux autonome et donc en dehors du champ d'action de la volonté. Il existe bien sûr des « trucs ». Tel adepte du yoga est capable d'agir sur son rythme cardiaque en contrôlant les muscles de la poitrine, mais cela équivaut à arrêter la circulation du sang dans une artère du poignet en appuyant dessus avec le pouce. On peut aussi accélérer les battements du cœur en se mettant dans un état d'anxiété, mais il s'agit là de la manipulation consciente du système nerveux autonome. Est-il possible d'accélérer le rythme cardiaque ou d'augmenter la tension simplement par la seule volonté, sans l'intervention de facteurs externes ?

Au début des années soixante, le psychologue américain Neal Elgar Miller et ses collègues ont fait des expériences sur le conditionnement, dans lesquelles on récompensait des rats quand, pour une raison ou pour une autre, leur tension augmentait, ou lorsque leur rythme cardiaque accélérait ou ralentissait. Finalement, et simplement pour le plaisir de recevoir une récompense, les rats avaient appris à provoquer volontairement une modification qui d'habitude était contrôlée par le système nerveux autonome – un peu comme s'ils avaient appris à appuyer sur un levier, et dans le même but.

Dans un programme expérimental sur des humains (de sexe masculin) que l'on récompensait par des éclairs de lumière révélant des photographies de femmes dénudées, on a pu démontrer que les volontaires pouvaient aussi répondre en augmentant ou en diminuant leur rythme cardiaque. Les volontaires ignoraient ce qu'on attendait d'eux pour déclencher l'éclair de lumière – et la vision des femmes nues – mais notaient seulement que plus le temps passait, plus ils voyaient les éclairs tant convoités.

Des expériences plus systématiques ont montré que si l'on renseigne constamment des sujets sur l'état d'une de leurs fonctions (ce dont ils n'ont généralement pas l'habitude) – comme par exemple la tension, le pouls, ou la température de la peau – ils sont capables, par un effort volontaire

(dont les modalités ne sont pas encore bien définies), d'en changer la valeur. Ce processus a pour nom le *biofeedback*.

On a fondé de grands espoirs à une époque sur les possibilités ouvertes par le biofeedback, pensant même qu'il permettrait d'accomplir avec une plus grande efficacité et beaucoup plus facilement les prouesses des mystiques indiens ; qu'il aiderait à contrôler ou améliorer des perturbations métaboliques sérieuses. Ces espoirs se sont quelque peu estompés au cours de ces dernières années.

L'HORLOGE BIOLOGIQUE

Il existe, dans les contrôles du système autonome, d'autres subtilités qui, jusqu'ici, étaient passées inaperçues. Comme les organismes vivants sont soumis à des rythmes naturels – les marées, l'alternance un peu plus lente du jour et de la nuit, les changements encore plus longs des saisons – il n'est pas étonnant qu'ils répondent de façon rythmique. Les arbres perdent leurs feuilles en automne et bourgeonnent au printemps ; les êtres humains tombent de sommeil le soir et se réveillent le matin.

On avait jusqu'ici complètement sous-estimé la complexité et la multiplicité des réponses rythmiques et leur nature automatique, qui persistent même en l'absence du rythme de l'environnement.

Ainsi, les feuilles des plantes s'élèvent et s'abaissent selon un rythme journalier qui suit les évolutions du Soleil. Ce phénomène s'observe facilement grâce aux photographies prises à intervalles réguliers. Ce cycle n'existe pas chez des plantules qu'on fait pousser dans le noir, mais la potentialité est là. Si on les expose à la lumière – une seule fois suffit – cette potentialité se transforme immédiatement en réalité. Le rythme commence, et continue même si la lumière est de nouveau retirée. Le rythme périodique varie un peu d'une plante à l'autre – de vingt-quatre à vingt-six heures en l'absence de lumière – mais il tourne toujours autour de vingt-quatre heures sous l'effet régulateur du Soleil. Un cycle de vingt heures peut être établi si l'on utilise une lumière artificielle dans des cycles de dix heures de jour et dix heures de nuit, mais dès que l'on éteint la lumière une fois pour toutes, le rythme de vingt-quatre heures se rétablit.

Ce rythme quotidien, sorte d'*horloge biologique*, qui fonctionne même en l'absence de tout indice extérieur, se retrouve dans toutes les formes de vie. Franz Halberg, de l'université du Minnesota, l'a nommé *rythme circadien*, du latin *circa dies*, qui signifie « environ un jour ».

L'être humain n'est évidemment pas à l'abri de tels rythmes. Des hommes et des femmes se sont portés volontaires pour vivre plusieurs mois d'affilée dans des grottes, sans aucun appareil pour mesurer le temps, et où ils ne savaient pas s'il faisait jour ou nuit. Dans ces conditions, les sujets perdent rapidement toute notion de temps, dorment et mangent de manière assez erratique. Au cours de ces expériences, ils notent leurs température, pouls, tension, ondes cérébrales, et les communiquent, ainsi que d'autres mesures, à la surface où des observateurs les corrèlent avec le temps. On s'est aperçu que, quel que fût l'état de confusion temporelle de ces troglodytes, leur rythme corporel, lui, fonctionnait parfaitement bien. Le rythme tournait obstinément autour d'une période de vingt-quatre heures, toutes les mesures augmentant et décroissant régulièrement, pendant tout le séjour au fond de la grotte.

Il ne s'agit pas ici simplement d'une question sans portée pratique. Dans la nature, la rotation de la Terre est constante, et le rythme alterné du jour et de la nuit est lui aussi constant ; l'homme ne peut rien y changer – à condition de rester au même endroit du globe ou de se déplacer, soit vers le nord, soit vers le sud. Si vous voyagez vers l'est ou vers l'ouest, et cela assez rapidement, vous modifiez l'heure de la journée. Il se peut que vous atterrissiez au Japon à l'heure du déjeuner (pour les Japonais) quand votre horloge biologique vous indique qu'il est temps d'aller vous coucher. A l'ère de l'avion à réaction, le voyageur rencontre souvent des difficultés à faire coïncider ses activités avec celles des autochtones autour de lui. S'il y parvient alors que son rythme de sécrétion hormonale, par exemple, ne coïncide pas avec le rythme de ses activités, il se fatigue et devient inefficace ; il souffre du *décalage horaire*.

Dans un autre ordre d'idées, la capacité qu'a l'organisme de résister à une certaine dose de rayons X ou d'assimiler certains types de médicaments dépend souvent de l'heure à laquelle est réglée l'horloge biologique. Un traitement médical devra par exemple varier suivant l'heure de la journée, ou, pour obtenir son effet maximum avec un minimum d'effets secondaires, il devra être administré à une heure bien définie.

Par quel mécanisme l'horloge biologique reste-t-elle aussi bien réglée ? Il semblerait que l'épiphyse (ou glande pinéale) soit destinée à remplir ce rôle (voir chapitre 15). Chez certains reptiles, la glande pinéale est particulièrement bien développée et sa structure ressemble à celle de l'œil. Chez le tuatara (nom mexicain du spenodon), sorte de lézard qui est la dernière espèce survivante dans son ordre et que l'on trouve seulement sur des petites îles près des côtes de Nouvelle-Zélande, l'*œil pinéal* est un morceau de peau qui recouvre le dessus du crâne, très proéminent pendant les six premiers mois de la vie, et dont on a démontré la sensibilité à la lumière.

L'épiphyse ne « voit » pas au sens habituel du terme, mais produit des substances dans des quantités croissantes et décroissantes de façon rythmique en réponse à la lumière ou à son absence. Il se peut donc qu'elle régule l'horloge biologique, même après que la lumière cesse d'être périodique (une sorte de conditionnement lui ayant appris sa leçon).

Mais alors, comment l'épiphyse fonctionne-t-elle chez les mammifères, où elle ne se trouve plus juste sous la peau au sommet du crâne, mais enfouie profondément au beau milieu du cerveau ? Existerait-il quelque chose de plus pénétrant que la lumière, quelque chose d'aussi rythmique dans la même acception du terme ? On spécule beaucoup actuellement sur l'action des rayons cosmiques. Ils possèdent un rythme circadien qui leur est propre, grâce au champ magnétique terrestre et au vent solaire, et peut-être cette force est-elle le régulateur externe recherché.

Même si l'on découvre la nature exacte de ce dernier, pourra-t-on jamais identifier précisément le mécanisme de l'horloge biologique ? Existe-t-il une quelconque réaction chimique dans notre corps qui suivrait un rythme circadien et contrôlerait tous les autres rythmes ? Y a-t-il une quelconque « réaction maîtresse » qui *seule* serait impliquée dans ce processus ? S'il en est ainsi, elle reste encore à découvrir.

SONDER LES PROFONDEURS DU COMPORTEMENT HUMAIN

Il est fort peu probable, néanmoins, qu'on arrivera jamais à confiner un phénomène aussi complexe que la vie dans un cadre totalement

déterministe. Il est facile de défendre une position déterministe dans un domaine comme la réplication des acides nucléiques, et pourtant là aussi, les facteurs de l'environnement introduisent des erreurs qui provoquent des mutations, et par voie de conséquence, l'évolution. Il est d'ailleurs également improbable que l'on pourra jamais prédire exactement le cours de l'évolution.

Nous savons grâce à la mécanique quantique qu'il existe des imprécisions et des incertitudes inhérentes à la nature des objets ; et que plus un objet est léger, moins il est massif, plus ces incertitudes sont grandes. Le comportement des électrons est d'une certaine manière imprévisible, et certains pensent que tant que nous serons incapables de les mesurer, il sera impossible de connaître toutes leurs propriétés. Il se pourrait même que l'état de l'Univers soit, d'une façon très subtile, influencé à chaque instant par les observations et les mesures faites par les hommes. (Cette idée porte le nom de *principe anthropique*, du mot grec signifiant « être humain ».)

On peut aisément concevoir qu'en certaines occasions, une décision ou un comportement humains (ou peut-être même un comportement chez des animaux inférieurs) dépende de la trajectoire indéterminée d'un électron quelque part dans le corps. En principe, voilà qui devrait démolir le déterminisme, mais sans pour autant d'ailleurs prouver le libre arbitre. Tout au plus cela introduit-il un facteur aléatoire, sans doute encore plus difficilement compréhensible que les deux autres cas.

Mais pas nécessairement plus difficile à interpréter. Les facteurs aléatoires peuvent être pris en compte s'ils sont assez nombreux. Les molécules de gaz se meuvent au hasard ; mais dans une quantité déterminée de gaz, il y a tellement de molécules que ces mouvements aléatoires s'annulent, et que des lois régissent avec une grande précision des propriétés telles que la température, la pression, et le volume.

Mais nous n'en sommes pas encore arrivés là dans l'étude du comportement humain ; je dirai même, au contraire, que certaines recherches dans ce domaine ont recours à des méthodes plus nébuleuses et plus ambiguës que le sujet qu'elles abordent.

Ces méthodes remontent à il y a environ deux siècles, à l'époque où un médecin autrichien, Franz Anton Mesmer, devint la coqueluche de l'Europe avec des expériences qui mettaient en jeu un outil puissant destiné à sonder les profondeurs du comportement humain. Il employa tout d'abord des aimants, puis simplement ses mains, obtenant des résultats par l'action de ce qu'il appelait le *magnétisme animal* (baptisé peu après *mesmérisme*) : la méthode consistait à plonger le patient dans un sommeil profond, puis à le déclarer guéri du mal dont il souffrait. Il est possible que Mesmer ait obtenu des guérisons (le pouvoir de la suggestion peut effectivement traiter certaines affections) et il ne fait aucun doute qu'il eut de nombreux et fervents disciples dont le marquis de La Fayette, encore tout enivré de son triomphe américain. Néanmoins, un comité scientifique, dont faisaient partie Lavoisier et Benjamin Franklin, enquêta avec un certain scepticisme certes, mais aussi avec impartialité, sur les activités de Mesmer, astrologue passionné et mystique convaincu ; le comité conclut à l'imposture et Mesmer dut prendre sa retraite au milieu du discrédit général.

Malgré tout, il avait lancé une mode. Dans les années 1850, un chirurgien anglais du nom de James Braid fit réapparaître l'*hypnotisme* (il

fut le premier à utiliser ce terme à la place de *mesmérisme*) comme traitement médical, et fut suivi par beaucoup d'autres médecins. Parmi eux, un docteur viennois qui s'appelait Josef Breuer commença à utiliser l'hypnose de manière systématique dans les années 1880 pour traiter les troubles de l'affectivité et les maladies mentales.

L'hypnotisme (du grec « endormir ») était bien sûr connu depuis l'Antiquité, et certains mystiques l'avaient souvent utilisé. Mais Breuer et les autres tentaient maintenant d'invoquer ses effets comme la preuve de l'existence d'un niveau *inconscient* de l'esprit. Des motivations inconnues de l'individu y étaient dissimulées, et l'hypnose pouvait les faire apparaître au grand jour. De là à penser que ces motivations disparaissaient du conscient parce qu'elles étaient associées à un sentiment de honte ou de culpabilité, et qu'elles étaient responsables de comportements irresponsables, irrationnels, voire vicieux, il n'y avait qu'un pas.

Breuer introduisit l'hypnose dans l'étude des causes cachées de l'hystérie, ainsi que d'autres troubles du comportement. Il avait à ses côtés un de ses élèves, Sigmund Freud. Pendant plusieurs années, ils traitèrent des patients ensemble, les plaçant sous hypnose légère et les encourageant à parler. Ils découvrirent que donner ainsi libre cours aux expériences et aux fantasmes profondément ensevelis dans leur inconscient produisait souvent un effet cathartique chez les patients, et soulageait leurs symptômes dès le réveil.

Freud conclut que pratiquement tous les souvenirs et toutes les motivations rejetés dans l'inconscient sont d'origine sexuelle. Les pulsions sexuelles réprimées par la société et par les parents sont enfouies dans les tréfonds de l'inconscient, mais cherchent toujours à s'exprimer en générant des conflits intenses qui sont d'autant plus dangereux qu'ils ne sont ni reconnus ni admis comme tels.

En 1894, après s'être brouillé avec son maître Breuer à cause d'un désaccord sur l'importance qu'il fallait accorder à la sexualité, Freud se lança seul dans le développement de ses idées sur les causes et le traitement des affections mentales. Il abandonna l'hypnose et encouragea ses clients à beaucoup parler – à dire tout ce qui leur passait par la tête. Au fur et à mesure que le patient prenait conscience que le médecin l'écoutait avec bienveillance et sans porter de jugement moral, lentement – parfois très lentement – l'individu commençait à se livrer, à se souvenir de choses depuis longtemps réprimées et oubliées. Freud appela cette longue analyse du *psyche* (du grec « âme », « esprit ») la *psychanalyse*.

Les travaux de Freud sur le symbolisme sexuel des rêves et ses descriptions de l'envie chez l'enfant de se substituer au parent du même sexe dans le lit conjugal (*complexe d'Œdipe* dans le cas des garçons, *complexe d'Électre* chez les filles – des noms empruntés à la mythologie grecque) horrifièrent les uns et fascinèrent les autres. Dans les années vingt, après le chaos dû à la Première Guerre mondiale, au beau milieu des traumatismes causés par la prohibition aux États-Unis et des bouleversements dans les mœurs aux quatre coins du monde, les idées de Freud ont touché une corde sensible, et la psychanalyse s'est pour ainsi dire hissée au rang de phénomène de masse.

Presqu'un siècle après ses débuts, la psychanalyse n'en demeure pas moins un art plus qu'une science. La rigueur de l'expérimentation, telle qu'elle existe en physique ou dans les autres sciences « pures et dures », est bien sûr un idéal extrêmement difficile à atteindre en psychiatrie. Pour

aboutir à leurs conclusions, les praticiens ne peuvent qu'avoir recours à leur intuition et à leur jugement, nécessairement subjectifs. La *psychiatrie* (dont la psychanalyse ne constitue qu'une des techniques) a sans aucun doute aidé plus d'un patient, mais elle n'a produit aucune cure extraordinaire, ni réduit de façon significative l'incidence des maladies mentales. Elle n'a pas non plus donné lieu à l'élaboration d'une théorie générale et universellement acceptée, comparable à la théorie microbienne dans les maladies infectieuses. C'est un fait qu'il y a presque autant d'écoles en psychiatrie qu'il y a de psychiatres.

Les maladies mentales graves se présentent sous plusieurs formes, de la dépression chronique à la fuite pure et simple de la réalité ; fuite dans un monde où certains détails ne correspondent pas à la façon dont la majorité d'entre nous conçoivent ou perçoivent les choses. On appelle généralement cette forme de psychose la *schizophrénie*, un terme inventé par le psychiatre suisse Eugen Bleuler. Ce terme recouvre une telle multitude d'affections qu'on ne peut plus l'appliquer aujourd'hui à une seule et même maladie. Le diagnostic de schizophrénie est rendu chez environ 60 % de tous les patients chroniques dans les hôpitaux psychiatriques américains.

Récemment encore, des traitements drastiques, tels que la lobotomie préfrontale ou les thérapies de choc, qui font appel à l'électricité ou à l'insuline (cette dernière technique a été introduite en 1933 par le psychiatre autrichien Manfred Sakel), étaient les seuls mis en œuvre. La psychiatrie et la psychanalyse ne sont pas d'un grand secours, sauf cas exceptionnel, quand le médecin peut encore, aux premiers stades de la maladie, communiquer avec le patient. Mais des recherches récentes sur certaines drogues et sur les processus chimiques qui se déroulent dans le cerveau (la *neurochimie*) permettent de nouveaux espoirs.

Même les Anciens savaient que la sève de certaines plantes peut provoquer des hallucinations (des sortes de mirages visuels, sonores, etc.) et que celle d'autres plantes peut induire des états euphoriques. Les prêtresses de Delphes, dans la Grèce antique, mâchaient des plantes avant de rendre leurs oracles. Des tribus indiennes dans le Sud-Ouest des États-Unis ont élevé au rang de rituel la mastication du peyotl (qui provoque des hallucinations colorées). Un des cas les plus célèbres est celui d'une secte musulmane qui vivait dans une forteresse au cœur des montagnes d'Iran et prenait du haschisch, le jus des feuilles de chanvre, plus connu sous le nom de *marijuana*. Cette drogue, prise au cours de cérémonies religieuses, donnait aux participants l'illusion qu'ils entrevoyaient le paradis vers lequel s'envolerait leur âme après la mort, et ils obéissaient aveuglément aux ordres de leur chef, le Vieil Homme des Montagnes, pour obtenir cette clé du paradis. Il ordonnait parfois le massacre de dirigeants ennemis et de fonctionnaires musulmans hostiles, donnant ainsi naissance au mot *assassin*, de *haschischin* (« personne qui prend du haschisch »). La secte terrorisa la région tout au long du XIIe siècle, jusqu'à ce que les hordes mongoles envahissent les montagnes en 1226, et tuent jusqu'au dernier assassin.

L'équivalent moderne des herbes euphoriques d'antan (mis à part l'alcool) est un ensemble de drogues connues sous le nom de *tranquillisants*. En fait, un de ces tranquillisants était déjà connu aux Indes depuis au moins l'an 1000 av. J.-C. sous la forme d'une plante appelée *Rauwolfia serpentinum*. C'est à partir des racines séchées de cette plante que des

chimistes américains ont extrait la *réserpine* en 1952, le premier des tranquillisants à la mode. De nombreuses substances produisant des effets similaires, mais avec des formules chimiques plus simples, ont été synthétisées depuis.

Les tranquillisants sont des sédatifs, mais avec une différence : ils réduisent l'anxiété sans affecter les autres activités mentales. Néanmoins, ils rendent les gens somnolents, et peuvent présenter d'autres effets secondaires. Dès leur apparition sur le marché, ils se sont révélés extrêmement utiles pour soulager et calmer les malades mentaux, même les schizophrènes. Les tranquillisants ne guérissent aucune maladie mentale, mais suppriment certains symptômes qui risqueraient de nuire à l'efficacité d'un traitement. En réduisant l'hostilité et le déchaînement de certains patients, en calmant leur crainte et leur anxiété, ils rendent moins nécessaire le recours aux contraintes physiques, facilitent l'établissement des contacts entre patient et psychiatre, et augmentent les chances qu'a le malade de sortir du milieu hospitalier.

Mais c'est surtout dans le grand public que les tranquillisants ont connu un immense succès ; ils sont considérés aujourd'hui par beaucoup comme une panacée qui guérit tous les maux.

LA DROGUE

La réserpine ressemble en fait de très près à une substance qui joue un grand rôle dans le cerveau. Une partie de sa molécule très complexe ressemble quelque peu à une substance qui s'appelle la *sérotonine*. La sérotonine a été découverte dans le sang en 1948, et a grandement intrigué les physiologistes depuis lors. On l'a trouvée dans la région de l'hypothalamus, mais elle est également présente ailleurs dans le cerveau et dans les tissus nerveux d'autres animaux, dont les invertébrés.

Qui plus est, on s'est aperçu que de nombreuses autres substances agissant sur le système nerveux central ressemblaient de près à la sérotonine. L'une d'elles est un produit qu'on trouve dans le venin de crapeau, la *bufoténine*. Une autre est la mescaline, la substance active dans le peyotl. La plus spectaculaire par ses effets s'appelle l'*acide lysergique diéthylamide* (plus connue sous le nom de LSD). En 1943, un chimiste suisse, Albert Hofmann, avala par erreur une faible quantité de ce produit dans son laboratoire, et se sentit soudain envahi par d'étranges sensations. En effet, ses perceptions n'avaient plus aucun rapport avec celles qu'il savait être la réalité de son environnement. Il était victime d'*hallucinations*, et le LSD est un exemple de ce que nous appelons aujourd'hui un *hallucinogène*.

Les adeptes du LSD disent qu'ils ressentent un « épanouissement mental » sans précédent sous l'effet de cette drogue – voulant sans doute dire par là qu'ils se sentent, ou croient se sentir plus en communion avec l'Univers que d'habitude. Mais tel est aussi le cas des ivrognes quand ils atteignent le stade du *delirium tremens*. La comparaison n'est pas aussi exagérée qu'il y paraît, car on a démontré qu'une petite dose de LSD, dans certains cas, peut provoquer les mêmes symptômes que la schizophrénie !

Mais quelles sont donc les causes de ces phénomènes ? Eh bien la sérotonine (qui ressemble structurellement au tryptophane, un acide aminé) peut être dégradée par une enzyme, l'*amine oxydase*, produite par

les cellules du cerveau. Supposons que l'action de cette enzyme soit inhibée par une substance concurrente de structure comparable à celle de la sérotonine – l'acide lysergique par exemple. Une fois l'enzyme de dégradation hors d'action, la sérotonine s'accumule dans les cellules du cerveau, et son niveau va atteindre un seuil critique. L'équilibre chimique est rompu dans le cerveau, provoquant un état schizophrénique.

Est-il possible que la schizophrénie naisse d'un tel déséquilibre induit naturellement dans le cerveau ? La nature héréditaire de cette affection laisse penser qu'un désordre métabolique (très probablement sous l'action d'un gène) est en cause. En 1962, on a découvert, dans l'urine de schizophrènes qui suivaient un traitement donné, une substance absente de l'urine de non-schizophrènes. Cette substance a maintenant été identifiée : il s'agit de la *diméthoxyphényléthylamine*, dont la structure se situe à mi-chemin entre l'adrénaline et la mescaline. Autrement dit, il semble que certains schizophrènes, par le fait d'une altération dans un processus métabolique, fabriquent leur propre substance hallucinogène, et sont ainsi « drogués » en permanence.

Personne ne réagit de la même façon à une dose donnée de telle ou telle drogue. Mais il est évidemment dangereux de jouer avec les mécanismes chimiques du cerveau. Aucune expérience, aussi « épanouissante » soit-elle, ne vaut vraiment la peine de courir le risque exhorbitant de la maladie mentale. Et pourtant, la réaction de la société vis-à-vis de la drogue – particulièrement de la marijuana, dont aucun effet nocif n'a encore été démontré à ce jour, contrairement à d'autres hallucinogènes – est pour le moins paradoxale. Combien d'entre nous s'insurgent contre l'utilisation de la drogue, alors qu'ils sont eux-mêmes complètement intoxiqués par l'alcool ou le tabac, deux fléaux responsables de méfaits considérables pour l'individu et la société ?

LA MÉMOIRE

La neurochimie porte également en elle l'espoir de mieux comprendre ce phénomène mental si insaisissable qu'est la *mémoire*. Il y aurait en fait deux sortes de mémoire : l'une à court terme, et l'autre à long terme. Quand vous regardez un numéro de téléphone dans l'annuaire, vous vous en souvenez sans trop de difficulté assez longtemps pour le composer sur votre cadran ; il est alors oublié automatiquement, et il y a de grandes chances que vous ne vous en souviendrez plus jamais. Mais un numéro que vous utilisez fréquemment entre dans le cadre de la mémoire à long terme. Même après plusieurs mois, vous allez sans doute parvenir à vous en souvenir.

Et pourtant, beaucoup de souvenirs qu'on aurait tendance à classer dans la catégorie de la mémoire à long terme sont perdus. Nous oublions énormément de choses, et certaines mêmes, hélas, vitales (comme l'étudiant qui planche pendant un examen le sait si bien). Mais sont-elles vraiment oubliées ? Se sont-elles véritablement évanouies, ou sont-elles simplement si bien stockées qu'il est difficile de les rappeler – enfouies, pour ainsi dire, sous le poids de détails superflus ?

L'« extraction » de ces souvenirs cachés est presque devenue une réalité aujourd'hui. Il est arrivé au chirurgien d'origine américaine Wilder Graves Penfield, au cours d'une opération qu'il pratiquait, sous anesthésie locale,

sur le cerveau d'un patient à l'université McGill de Montréal, de toucher accidentellement un point bien particulier, ce qui a provoqué chez le patient l'audition d'un air de musique. Ce phénomène s'est renouvelé plusieurs fois. Le patient pouvait revivre tous les détails de la sensation, tout en restant tout à fait conscient du présent. La stimulation d'une zone spécifique peut donc apparemment provoquer l'évocation de souvenirs très précis. La zone en question s'appelle le *cortex interprétatif.* Il est possible qu'un stimulus accidentel de cette partie du cortex provoque le phénomène de *déjà vu* ainsi que d'autres manifestations de *perception extra-sensorielle.*

Mais si la mémoire retient tant de détails, comment le cerveau leur fait-il assez de place ? On estime qu'au cours d'une vie, le cerveau peut emmagasiner un million de milliards d'unités d'information. Pour stocker une telle quantité de données, les unités de stockage doivent avoir la taille d'une molécule. Sinon, il n'y aurait de la place pour rien d'autre.

Les recherches se tournent actuellement vers un rôle potentiel de l'acide ribonucléique (ARN) dont les nerfs, assez curieusement, sont très riches, plus riches que pratiquement n'importe quel autre type de cellules dans le corps. Cela est curieux, car l'ARN est impliqué dans la synthèse des protéines (voir chapitre 13), et on le trouve donc en grande quantité dans les tissus qui produisent eux-mêmes de grandes quantités de protéines, soit parce qu'ils sont dans une phase de croissance aiguë, soit parce qu'ils produisent d'importantes quantités de sécrétions riches en protéines. Les cellules nerveuses n'entrent ni dans la première, ni dans la seconde catégorie.

Un neurologue suédois, Holger Hyden, a mis au point des techniques permettant d'isoler des cellules individuelles du cerveau, puis d'en déterminer le contenu en ARN. Il a pris des rats qui venaient de subir un apprentissage – par exemple, se tenir en équilibre sur un fil pendant de longues périodes. Il a découvert en 1959 que leurs cellules cérébrales contenaient jusqu'à 12 % d'ARN de plus que celles des rats qui n'avaient subi aucun apprentissage.

La molécule d'ARN est si grande et si complexe que si chaque unité de mémoire stockée est effectivement caractérisée par un agencement spécifique des bases de cette molécule, il n'y a plus de souci à se faire sur la question de la capacité de la mémoire : l'ARN comporte tellement de possibilités d'agencement différentes que le nombre un million de milliards est insignifiant en comparaison.

Mais doit-on se limiter à l'étude de l'ARN seul ? Les molécules d'ARN sont fabriquées en suivant exactement la séquence des molécules d'ADN dans les chromosomes. Est-il possible que chacun d'entre nous possède des provisions quasi illimitées de souvenirs potentiels – une *banque de souvenirs,* en quelque sorte – dans les molécules d'ADN que nous avons depuis la naissance, et qu'elles soient d'une façon ou d'une autre appelées et activées par des événements réels, avec les modifications voulues ?

Et n'y aurait-il que l'ARN d'impliqué ? La fonction essentielle de l'ARN est de fabriquer des molécules de protéines spécifiques. Serait-ce la protéine, plutôt que l'ARN, qui serait en cause dans le processus de la mémoire ?

Une façon de vérifier cette hypothèse consiste à utiliser une substance, la puromycine, qui interfère avec la fabrication de la protéine en agissant sur l'ARN. L'équipe américaine constituée de Louis Barkhouse Flexner et sa femme Josepha Barbara Flexner a appris à des souris à sortir d'un

labyrinthe, puis leur a injecté, immédiatement après, de la puromycine. Les souris ont oublié ce qu'elles avaient appris. La molécule d'ARN était toujours présente, mais la molécule clé de la protéine ne pouvait plus être fabriquée. Par l'utilisation de la puromycine, les Flexner ont montré que bien que l'on puisse effacer de cette manière la mémoire à court terme chez le rat, il était impossible d'obtenir le même résultat avec la mémoire à long terme. Il est probable que les protéines impliquées dans cette dernière avaient déjà été fabriquées.

Pourtant, il est possible que la mémoire soit un phénomène bien plus complexe, et que la simple explication moléculaire soit très insuffisante. On pense par exemple que différents processus impliqués dans l'activité nerveuse entrent également en jeu. Il reste décidément encore beaucoup à faire dans ce domaine.

Les automates

Ce n'est que très récemment que les phares de la science se sont braqués de toute leur puissance sur le fonctionnement des tissus et des organes vivants, afin que ces processus – qui se sont développés vaille que vaille au cours des quelques milliards d'années d'évolution – puissent être reproduits par des machines fabriquées par l'homme. Cette étude s'appelle la *bionique* (en anglais *bionics*, mot tiré de l'expression « *bio*logical electro*nics* », mais pris dans un sens beaucoup plus large), un terme inventé par l'ingénieur américain Jack Steele en 1960.

Pour mieux apprécier ce que la bionique peut réaliser, prenons l'exemple de la structure de la peau du dauphin. Ce dernier nage à des vitesses qui nécessiteraient une dépense d'énergie égale à 2,6 chevaux si les turbulences de l'eau qui l'entoure étaient les mêmes que celles créées par un bateau de taille identique. Pour une raison inconnue, l'eau s'écoule le long du corps du dauphin sans turbulence ; la dépense d'énergie nécessaire pour surmonter la résistance de l'eau est donc pratiquement nulle. Apparemment, l'explication tient à la nature de sa peau. Si l'on pouvait reproduire cet effet et l'adapter à la coque d'un navire, on pourrait – simultanément – augmenter la vitesse d'un paquebot et diminuer sa consommation de carburant.

Autre exemple : le biophysicien américain Jerome Lettvin a étudié en détail la rétine de la grenouille en insérant de minuscules électrodes de platine dans le nerf optique. Il a trouvé que la rétine ne se limite pas à transmettre au cerveau un mélange de signaux lumineux et noirs et à lui laisser tout le travail d'interprétation. En fait, la rétine est constituée de cinq types différents de cellules, dont chacune assume un rôle propre. Un type de cellule réagit aux transitions – c'est-à-dire aux modifications soudaines de luminosité, comme dans le cas d'un arbre à contre-jour sur un fond de ciel ensoleillé. Un autre type est sensible aux objets sombres de forme arrondie (les insectes que mange la grenouille). Un troisième réagit à tout ce qui bouge très rapidement (une créature dangereuse qu'il est préférable d'éviter). Le quatrième réagit à la lumière tombante ; et le cinquième, au bleu de l'eau d'un étang. Autrement dit, le message rétinien

parvient au cerveau déjà considérablement analysé. Si l'homme pouvait construire des détecteurs basés sur les mêmes principes que ceux de la rétine de la grenouille, ces détecteurs seraient plus sensibles et plus souples que ceux existant actuellement.

Mais si l'on veut réellement construire une machine qui imite un mécanisme vivant, pourquoi ne pas imiter tout de suite ce mécanisme extraordinaire et unique qui nous intéresse tant – le cerveau humain ?

L'esprit humain n'est pas qu'une « simple » machine, loin de là. Cependant l'esprit, sans aucun doute le phénomène le plus complexe que nous connaissions, comporte des caractéristiques qui rappellent plus ou moins celles d'une machine. Et les ressemblances peuvent être importantes.

Donc, si l'on analyse ce qui distingue l'esprit humain d'autres esprits (sans parler des objets matériels), une chose est frappante : plus que tout autre objet ou phénomène, vivant ou non vivant, l'esprit humain est un système autorégulé. Il est capable non seulement de se contrôler, mais aussi de contrôler son environnement. Il s'adapte aux changements de l'environnement, non pas en cédant du terrain, mais en réagissant selon ses propres désirs et ses propres règles. Voyons à quel point une machine peut approcher cet idéal.

La valve représente sans doute la forme la plus simple de mécanisme autorégulé. Des versions grossières ont été mises au point dès 50 apr. J.-C. par Héron d'Alexandrie, qui les utilisait dans un appareil distribuant des liquides automatiquement. La marmite autoclave inventée par Denis Papin en 1679 est une version très élémentaire de la soupape de sûreté. Pour que le couvercle reste sur la marmite malgré la pression, il plaça un poids dessus, mais suffisamment léger pour que le couvercle puisse se soulever avant que la pression ne fasse exploser le tout. La cocotte-minute ou la chaudière à vapeur d'aujourd'hui utilisent des moyens plus sophistiqués dans le même but (une bonde, par exemple, qui fond quand la température est trop élevée) ; mais le principe est le même.

LE FEED-BACK

Bien sûr, il s'agissait là d'un type de régulation qui ne fonctionne qu'une fois. Mais on peut facilement imaginer des exemples de régulation continue. En 1745, un Anglais, Edmund Lee, a fait breveter un mécanisme simple permettant à un moulin à vent de toujours rester face au vent. Il avait inventé un dispositif muni de petites ailettes en queue d'hirondelle dans lequel s'engouffrait le vent lorsqu'il changeait de direction ; ces ailettes actionnaient une série d'engrenages qui faisaient tourner l'ensemble du moulin à vent, permettant ainsi à ses grandes ailes de rester perpendiculaires au vent. Dans cette nouvelle position, les petites ailettes ne tournaient plus ; elles ne tournaient que lorsque le moulin n'était plus face au vent.

Mais l'archétype des autorégulateurs mécaniques modernes est le *régulateur de Watt,* du nom de son inventeur James Watt, qui le mit au point pour son moteur à vapeur (figure 17.4). Pour maintenir un débit constant de vapeur dans son moteur, Watt imagina un système fait d'un arbre vertical auquel deux poids sont attachés latéralement à des tiges munies de charnières, ce qui permet aux poids de se déplacer de bas en

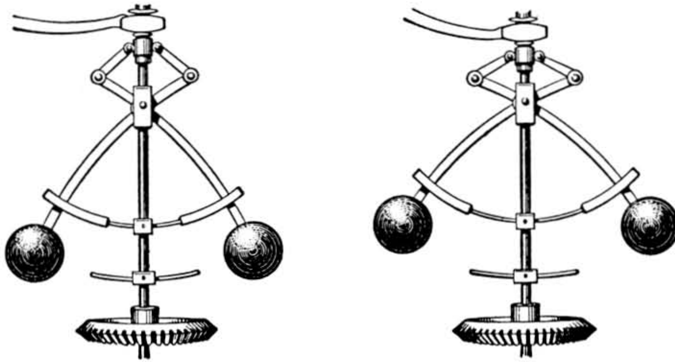


Figure 17.4. Le régulateur de Watt.

haut. La vapeur fait tourner l'arbre. Lorsque la pression augmente, l'arbre tourne plus vite, la force centrifuge fait monter les poids, ce qui provoque la fermeture partielle d'une valve, et diminue l'arrivée de vapeur. La pression baisse, l'arbre tourne donc moins vite, et la pesanteur ramène les poids vers le bas, ce qui provoque l'ouverture de la valve. Le régulateur maintient ainsi la vitesse de l'arbre, et par conséquent le débit de vapeur, à un niveau constant. Chaque écart entraîne un processus qui corrige la déviation. Cela s'appelle le *feed-back* : c'est l'erreur qui envoie continuellement des informations servant à mesurer le niveau de correction requis.

Un exemple bien connu d'appareil basé sur la notion de feed-back est le *thermostat*, tout d'abord utilisé sous une forme rudimentaire par l'inventeur néerlandais Cornelis Drebble au début du XVIIe siècle. Une version plus sophistiquée, utilisée encore de nos jours, fut inventée, dans son principe, par un chimiste écossais du nom de Andrew Ure en 1830. Sa partie principale est constituée de deux morceaux de métaux différents, juxtaposés et soudés. Les deux métaux se dilatent et se contractent différemment sous l'action de la chaleur, provoquant une courbure. Admettons que le thermostat soit réglé sur 20 °C. Si la température de la pièce descend au-dessous de cette valeur, le thermocouple se plie de telle façon qu'il fait contact et ferme un circuit électrique qui allume le système de chauffage. Lorsque la température monte au-dessus de 20 °C, le thermocouple se plie de l'autre côté et coupe le contact. De cette façon, la chaudière régule son propre fonctionnement grâce à un feed-back.

C'est aussi le feed-back qui contrôle le fonctionnement du corps humain. Pour prendre un exemple parmi tant d'autres, le taux de glucose dans le sang est sous le contrôle du pancréas, la glande qui sécrète l'insuline, un peu comme la température d'une maison est contrôlée par la chaudière. De la même manière que le fonctionnement de la chaudière est régulé par un écart de la température par rapport à la norme, la sécrétion de l'insuline est régulée par l'écart de la concentration en glucose par rapport à la norme. Un niveau de glucose trop élevé « allume » l'insuline, tout comme une température trop basse allume la chaudière. Et de la même façon qu'on peut régler un thermostat à une température plus élevée, un changement interne du corps, comme la sécrétion d'adrénaline, peut

provoquer l'établissement d'une nouvelle norme, à laquelle se pliera le pancréas.

Le processus d'autorégulation chez les organismes vivants, qui leur permet de maintenir certaines normes à des niveaux constants, a été baptisé *homéostasie* par le physiologiste américain Walter Bradford Cannon, qui a étudié ce phénomène dans les premières décades du XXe siècle.

Le processus de feed-back dans les systèmes vivants est essentiellement le même que celui des machines et on ne lui donne habituellement aucun nom particulier. Le terme *biofeedback* est une convention un peu artificielle utilisée dans les cas où le contrôle volontaire des fonctions nerveuses autonomes entre en jeu.

Dans la plupart des systèmes, vivants ou non vivants, il existe un certain retard dans la réponse au feed-back. Par exemple, une chaudière, une fois éteinte, continue pendant quelque temps à émettre une chaleur résiduelle ; inversement, quand on l'allume, il lui faut un certain temps avant de chauffer. Par conséquent, la température de la pièce ne se maintient pas à 20 °C mais oscille autour de cette valeur ; elle la dépasse constamment d'un côté ou de l'autre. Ce phénomène, la *chasse*, a été tout d'abord étudié dans les années 1830 par George Airy, astronome de Sa Majesté la reine d'Angleterre, alors qu'il mettait au point des systèmes destinés à asservir ses télescopes au déplacement de la Terre.

L'irrégularité est caractéristique de la plupart des processus vivants, du contrôle du niveau de glucose dans le sang à la conduite consciente. Quand on tend la main pour saisir un objet, l'approche ne se fait pas d'un trait mais par une série de mouvements dont la vitesse et la direction s'ajustent continuellement, grâce aux muscles qui corrigent les écarts par rapport à la ligne désirée, ces écarts étant mesurés par l'œil. Les corrections sont tellement automatiques qu'elles passent inaperçues. Mais observez un enfant qui n'a pas encore maîtrisé le processus du feed-back visuel en train d'attraper un objet : il surestime ou sous-estime la distance, parce que les corrections apportées par ses muscles sont encore imprécises. Et les personnes atteintes de lésions nerveuses qui interfèrent avec ce processus sont victimes d'une incoordination musculaire, provoquant des oscillations désordonnées.

La main normale, expérimentée, se dirige sans à-coups vers sa cible et s'arrête en temps voulu parce que le centre de contrôle sait prévoir les corrections. Ainsi, lorsque vous conduisez votre voiture dans un tournant, vous commencez à relâcher le volant avant d'avoir complètement négocié le virage, afin que les roues se retrouvent bien parallèles à la route une fois arrivées dans la ligne droite. Autrement dit, la correction est appliquée à temps pour éviter une trop grande erreur d'estimation.

C'est au cervelet que revient apparemment le soin d'ajuster le mouvement grâce au feed-back. Il prévoit ce que sera la position du bras quelques instants plus tard, et « organise » le mouvement en conséquence. C'est lui qui fait varier constamment le tonus des grands muscles du torse, ceux qui maintiennent notre équilibre et nous font nous tenir droit quand nous sommes debout. C'est une tâche difficile que de se tenir debout et « ne rien faire » ; nous savons tous à quel point cette position est fatigante.

Eh bien, ce principe peut être appliqué à une machine. Il est possible de concevoir un système dans lequel, lorsque la condition voulue est presque atteinte, la marge entre l'état effectif et l'état désiré provoque, à un seuil prédéterminé, la coupure automatique de la force correctrice avant

toute surcompensation. En 1868, l'ingénieur français Léon Farcot a mis en œuvre ce principe en inventant un système de contrôle automatique de gouvernail à vapeur. Quand le gouvernail approche la position voulue, le système ferme automatiquement une valve ; lorsqu'il atteint la bonne position, la pression de la vapeur tombe à zéro. Si le gouvernail s'écarte de cette position, sa course provoque l'ouverture de la valve et la pression le pousse dans la bonne direction. Farcot a baptisé *servomécanisme* son système qui a en quelque sorte inauguré l'ère de l'*automation* (terme inventé par l'ingénieur américain John Diebold).

LES AUTOMATES D'ANTAN

L'invention de mécanismes imitant la faculté de prévoir et le jugement de l'homme, même grossièrement, n'a pas manqué de susciter chez certains l'idée de machines qui reproduiraient ses faits et gestes de manière plus ou moins achevée, des *automates*. Mythes et légendes en sont remplis.

Mais copier ainsi les prouesses des dieux et des magiciens a tout d'abord nécessité le développement progressif des horloges au cours du Moyen Age. Grâce à des *mécanismes d'horlogerie* de plus en plus compliqués, c'est-à-dire des combinaisons subtiles de rouages actionnant certains mouvements, selon une séquence déterminée et au moment choisi, on a pu envisager la fabrication d'objets qui imitent la vie de façon saisissante.

Le XVIIIe siècle a été l'âge d'or des automates. Le dauphin du roi de France faisait construire des petits soldats automates ; un prince indien possédait un tigre mécanique grandeur nature.

Cependant ces caprices princiers furent vite dépassés par des entreprises plus mercantiles. En 1738, le Français Jacques de Vaucanson construisit un canard mécanique en cuivre qui cancanait, se baignait, buvait de l'eau, mangeait des graines, semblait les digérer, puis les excréter. On payait pour voir le canard, et il rapporta beaucoup d'argent à son propriétaire ; hélas, il n'existe plus.

Un automate fabriqué postérieurement a survécu, lui, et se trouve aujourd'hui dans un musée de Neuchâtel, en Suisse. C'est un automate scribe que l'on doit au génie de Pierre Jacquet-Droz qui le fabriqua en 1774. Il se présente sous la forme d'un jeune garçon qui trempe sa plume dans un encrier et écrit une lettre.

Il va sans dire que de tels automates ne sont doués d'aucune flexibilité. Ils peuvent *seulement* reproduire les mouvements qui leur sont dictés par leur mécanique interne.

Mais il ne fallut pas attendre longtemps avant que les principes de l'automatisme s'assouplissent et se tournent vers des applications plus utiles, bien que moins spectaculaires.

Dès 1801, le tisserand français Joseph Marie Jacquard, avec son *métier à tisser*, en apporte le premier exemple probant.

Dans ce type de métier, les aiguilles passent habituellement dans des trous pratiqués dans des blocs de bois, où elles s'accrochent aux fils pour former les diverses interconnexions du tissage. Supposons maintenant qu'on interpose une carte perforée entre les aiguilles et les fils ; les trous de la carte laissent passer les aiguilles çà et là, afin qu'elles accomplissent leur tâche comme avant. Aux endroits où il n'y a pas de trous, les aiguilles

sont arrêtées. De cette façon, le système permet à certaines interconnexions de s'établir, et pas à d'autres.

Si nous imaginons maintenant plusieurs cartes perforées, avec des agencements de trous différents, et qu'on les insère dans la machine selon un ordre déterminé, nous obtiendrons une variété de points qui produiront un motif. En faisant varier l'ordre des cartes, n'importe quel motif, en principe, peut être tissé automatiquement. Aujourd'hui, nous dirions que les cartes servent à *programmer* le métier à tisser, qui fabrique alors un produit, apparemment sans aide extérieure, comparable à une création artistique.

L'aspect important du métier de Jacquard, c'est qu'il doit son succès considérable (il y en avait déjà onze mille en France en 1812, et il s'est répandu dans toute l'Angleterre après les guerres napoléoniennes) à une simple dichotomie oui/non. Un trou existe ou il n'existe pas à un endroit particulier, et rien d'autre n'est nécessaire que la suite de oui-non-oui-oui-non qui figure sur la carte.

Depuis lors, un nombre croissant de dispositifs compliqués destinés à imiter la pensée humaine se sont inspirés de méthodes de plus en plus subtiles permettant de jouer sur cette dichotomie. En fait, sa base mathématique avait été démontrée au XVIIe siècle, après des milliers d'années d'efforts visant à mécaniser les calculs arithmétiques et à aider (de plus en plus efficacement) nos opérations mentales.

LE CALCUL ARITHMÉTIQUE

Il fait peu de doute que les premiers outils ayant servi à cet usage ont été les doigts. Les mathématiques sont nées le jour où des êtres humains ont utilisé leurs doigts pour représenter des chiffres, seuls ou en combinaison. Ce n'est pas sans raison que l'adjectif *digital* se rapporte à la fois à « doigt » et à « chiffre ».

A partir de là, l'étape suivante a sans doute consisté à remplacer les doigts par d'autres objets – peut-être des petits cailloux. Il y a plus de cailloux que de doigts, et les cailloux permettent de conserver les résultats intermédiaires pour un usage ultérieur. De nouveau, ce n'est pas sans raison que le mot *calcul* vient du latin « caillou ».

Des cailloux ou des boules, les uns alignés dans les rainures, les autres enfilées, constituent le boulier (ou abaque), le premier outil mathématique à offrir véritablement une souplesse d'emploi (figure 17.5). Grâce à lui, il est devenu facile de représenter les unités, les dizaines, les centaines, les milliers, etc. En manipulant les cailloux ou les boules d'un abaque, on peut rapidement effectuer une addition telle que $576 + 289$. De plus, n'importe quel instrument capable d'additionner peut également multiplier, puisque la multiplication n'est rien d'autre qu'une suite d'additions. Et la multiplication rend possible l'élévation à la puissance, puisque cette dernière n'est rien d'autre qu'une suite de multiplications (par exemple, 4^5 est une forme de notation sténographique qui signifie $4 \times 4 \times 4 \times 4 \times 4$). Enfin, en prenant l'instrument à l'envers, pour ainsi dire, on peut effectuer les opérations de soustraction, de division et d'extraction de racine.

On peut dire que le boulier est le second *ordinateur digital* (le premier, bien entendu, étant les doigts). Pendant des milliers d'années, le boulier

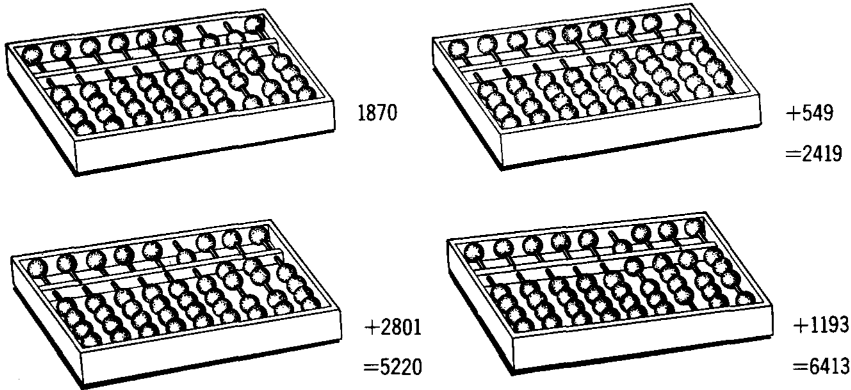


Figure 17.5. Additionner avec un boulier. Chaque boule sous la barre vaut 1 ; chaque boule au-dessus de la barre vaut 5. Une boule prend sa valeur quand elle est poussée vers la barre. Ainsi, en haut de la figure à gauche, la colonne de droite vaut 0 ; celle à sa gauche vaut 7 ou (5 + 2) ; celle plus à gauche, 8 ou (5 + 3) ; encore plus à gauche, elle vaut 1 : le nombre indiqué est donc 1870. Si l'on veut ajouter 549 à ce nombre, la colonne de droite devient 9 ou (9 + 0) ; l'addition suivante (4 + 7) devient 1 plus 1 de retenue, ce qui veut dire qu'une boule est poussée vers la barre dans la colonne suivante ; la troisième addition est 9 + 5, ou 4 plus 1 de retenue ; et la quatrième addition est 1 + 1 ou 2 ; l'addition donne 2419, comme le montre le boulier. La manœuvre simple de la retenue permet d'effectuer des calculs très rapidement ; un habitué peut additionner ainsi plus vite qu'avec une machine à calculer, comme l'a démontré un test en 1946.

est resté la forme de calculateur la plus évoluée. Il a presque disparu d'Europe après la chute de l'Empire romain, et a connu un regain de popularité quand le pape Sylvestre II l'a réintroduit vers l'an 1000, probablement de l'Espagne maure où son utilisation s'était poursuivie. A son retour, il fut accueilli comme une nouveauté orientale, car le souvenir de son héritage occidental s'était perdu dans la nuit des temps.

L'abaque n'a pas été remplacé avant l'introduction d'une notation numérique imitant son fonctionnement. (Cette notation, que nous connaissons aujourd'hui familièrement sous le nom de *chiffres arabes*, trouve son origine aux Indes vers 800 apr. J.-C., et fut introduite en Occident vers l'an 1200 par le mathématicien Léonard de Pise.)

Dans cette nouvelle notation, les neuf cailloux de la colonne des unités du boulier sont représentés par neuf symboles, et ces mêmes symboles sont utilisés pour la colonne des dizaines, celle des centaines, et celle des milliers. Les boules qui diffèrent seulement par leur position sont donc remplacées par des symboles qui eux aussi diffèrent par leur position, si bien que dans le nombre 222, par exemple, le premier 2 représente 200, le second 20, et le troisième représente le chiffre 2 ; on a ainsi 200 + 20 + 2 = 222.

Cette « notation positionnelle » n'est possible que si l'on reconnaît l'importance d'un fait qu'avaient négligé les utilisateurs du boulier dans

les temps anciens. Bien qu'il n'y ait que neuf boules dans chaque colonne de l'abaque, il existe en fait dix positions possibles. On peut effectivement utiliser non seulement n'importe quel nombre de boules, de une à neuf, dans une colonne donnée, mais on peut aussi n'en utiliser *aucune*, c'est-à-dire donner une signification nulle à la place vide. Ceci avait totalement échappé aux grands mathématiciens grecs et ne fut reconnu qu'au IXe siècle, lorsqu'un Hindou resté anonyme eut l'idée de représenter la dixième alternative par un symbole spécial que les Arabes ont appelé « sifr » (« vide ») et qui a donné « chiffre », et dans une forme plus corrompue, « zéro ».

Un autre outil puissant s'est dégagé de l'utilisation de l'exposant pour exprimer les puissances d'un nombre. Exprimer 100 sous la forme 10^2, 1 000 sous la forme 10^3, 100 000 sous la forme 10^5, etc., est très pratique pour de nombreuses raisons ; non seulement il devient plus facile d'écrire les grands nombres, mais la multiplication et la division se réduisent à de simples additions et soustractions d'exposants (par exemple, $10^2 \times 10^3 = 10^5$) ; et l'élévation à la puissance ou l'extraction de racine se réduisent à un problème de multiplication et de division d'exposants (par exemple, la racine cubique de 1 000 000 000 est $10^{6/3} = 10^2$). Bon, tout cela est très bien, mais il existe très peu de nombres qui peuvent être exprimés sous une forme exponentielle simple. Que peut-on faire d'un nombre comme 111, par exemple ? La réponse à cette question se trouve être les tables de logarithmes.

Le premier à s'intéresser à ce problème fut, au XVIIe siècle, le mathématicien écossais John Napier. Évidemment, exprimer un nombre tel que 111 sous la forme d'une puissance de 10 nécessite l'utilisation d'un exposant fractionnaire de 10 (l'exposant se situe entre 2 et 3). En règle générale, l'exposant sera fractionnaire quand le nombre en question n'est pas un multiple du nombre de base. Napier mit au point une méthode pour calculer les exposants fractionnaires de nombres, et il appela ces exposants des *logarithmes*. Peu après, le mathématicien anglais Henry Briggs simplifia cette technique et mit au point des logarithmes de base 10. Les *logarithmes de Briggs* sont moins commodes en calcul infinitésimal, mais ils sont très utilisés pour les opérations courantes.

Tout exposant fractionnaire est irrationnel, c'est-à-dire qu'il ne peut être exprimé sous la forme d'une fraction ordinaire. Il ne peut être exprimé que sous la forme d'un nombre comportant une infinité de décimales sans séquence répétitive. Mais on peut calculer un tel nombre avec autant de décimales que l'exige la précision voulue.

Par exemple, admettons que nous voulions multiplier 111 par 254. Le logarithme briggsien de 111 à cinq décimales près est 2,04532, et celui de 254 est 2,40483. L'addition de ces logarithmes nous donne $10^{2,04532} \times 10^{2,40483} = 10^{4,45015}$. Ce nombre donne approximativement 28 194, qui est effectivement le produit de 111×254. Si nous voulons être encore plus précis, il nous est possible d'utiliser des logarithmes à six décimales près ou plus.

Les tables de logarithmes ont énormément simplifié les calculs. En 1622, le mathématicien anglais William Oughtred facilita encore plus les choses en inventant la *règle à calcul*. Il s'agit en fait de deux règles sur lesquelles est inscrite une *échelle logarithmique*, dont la distance entre les nombres décroît proportionnellement à la grandeur de ces derniers ; par exemple, la première division comprend les nombres entre 1 et 10 ; la seconde

division, de même longueur, comprend les nombres de 10 à 100 ; la troisième les nombres de 100 à 1 000, etc. En faisant glisser l'une des règles contre l'autre jusqu'à la position voulue, il est possible de lire le résultat d'une opération impliquant la multiplication ou la division. La règle rend les calculs aussi faciles que l'addition ou la soustraction à l'aide du boulier ; mais dans les deux cas, bien sûr, une certaine expérience de l'instrument est nécessaire.

LES MACHINES À CALCULER

C'est au mathématicien français Blaise Pascal que l'on doit, en 1642, l'invention de la première machine à calculer véritablement automatique. Dans cette dernière, point n'est besoin de déplacer des boules dans chaque colonne comme c'est le cas pour l'abaque. La machine est faite d'une série de roues et d'engrenages. Si on avance la première roue – celle des unités – de dix crans à sa position 0, elle fait tourner la seconde d'un cran à sa position 1 ; ensemble, les deux roues indiquent le nombre 10. Lorsque la roue des dizaines atteint son 0, la troisième tourne d'un cran, pour donner 100, etc. (C'est sur ce même principe que fonctionne le compteur kilométrique d'une automobile.) On pense que Pascal fit fabriquer une cinquantaine de ces machines, dont cinq existent encore aujourd'hui.

La machine à calculer de Pascal pouvait faire des additions et des soustractions. En 1674, le mathématicien allemand Gottfried Wilhelm von Leibnitz apporta des innovations dans le système des roues et des engrenages, qui rendirent la multiplication et la division aussi aisées que les deux autres opérations. En 1850, un inventeur américain, D.D. Parmalee, breveta une invention qui allait entraîner une véritable révolution en facilitant considérablement l'emploi des machines à calculer : au lieu de manipuler directement les roues, il introduisit une série de clés – en appuyant du doigt sur une clé donnée, les roues tournaient automatiquement pour afficher le nombre voulu. C'est le mécanisme utilisé dans la bonne vieille caisse enregistreuse.

Mais Leibnitz ne s'est pas arrêté là. Les efforts qu'il a prodigués pour mécaniser le calcul lui ont permis de concevoir sa simplification ultime, le système binaire.

Les êtres humains ont pris pour habitude d'utiliser un système basé sur le nombre 10 (le système *décimal*), dans lequel dix chiffres différents (0, 1, 2, 3, 4, 5, 6, 7, 8, 9) représentent, dans des quantités et des combinaisons variées, tous les nombres imaginables. Dans certaines cultures, on utilise d'autres bases (il existe des systèmes à base cinq, base vingt, base douze, base soixante, etc.), mais la base dix est de loin la plus répandue. Le fait que nous ayons dix doigts n'est sans doute pas étranger à l'affaire.

Leibnitz a remarqué que *n'importe quel nombre* pouvait servir de base, et que, à bien des égards, le plus simple à adapter à une machine était le système à base deux (ou système *binaire*).

La notation binaire n'a recours qu'à deux chiffres : 0 et 1. Elle exprime tous les nombres en termes de puissance de 2. Ainsi, le chiffre 1 est représenté par 2^0, le chiffre 2 par 2^1, 3 par $2^1 + 2^0$, 4 par 2^2, etc. Comme dans le système décimal, la puissance est indiquée par la position du

symbole. Par exemple, le chiffre 4 est représenté par 100, lu de la façon suivante : $(1 \times 2^2) + (0 \times 2^1) + (0 \times 2^0)$, ou $4 + 0 + 0 = 4$ dans le système décimal.

Prenons comme exemple le nombre 6 413. Dans le système décimal, on peut l'écrire $(6 \times 10^3) + (4 \times 10^2) + (1 \times 10^1) + (3 \times 10^0)$; il faut se souvenir que n'importe quel nombre élevé à la puissance 0 vaut 1. Dans le système binaire, donc, pour former un nombre, on ajoute des chiffres multipliés par des puissances de 2, au lieu de puissances de 10. La puissance la plus élevée de 2 qui donne un résultat inférieur à 6 413 est 12 ; 2^{12} est égal à 4 096. Si nous ajoutons 2^{11}, ou 2 048, cela donne 6 144, soit 269 de moins que 6 413. Maintenant, 2^8 ajoute 256 de plus, et il manque encore 13 ; on peut alors ajouter 2^3, soit 8, et il manque 5 ; puis 2^2, soit 4, et reste 1 ; et 2^0 est égal à 1. Donc, on peut écrire le nombre 6 413 comme suit :

$(1 \times 2^{12}) + (1 \times 2^{11}) + (1 \times 2^8) + (1 \times 2^3) + (1 \times 2^2) + (1 \times 2^0)$.

Mais comme dans le système décimal, chaque chiffre dans un nombre, lu de gauche à droite, doit représenter la puissance voisine la plus petite. De la même manière que dans le système décimal, on représente les additions de la troisième, seconde, première et zéroième puissance de 10 quand on écrit le nombre 6 413, on doit dans le système binaire représenter les additions des puissances de 2, de 12 jusqu'à 0. Sous forme de tableau, on a ainsi :

$1 \times 2^{12} =$	4 096
$1 \times 2^{11} =$	2 048
$0 \times 2^{10} =$	0
$0 \times 2^9 =$	0
$1 \times 2^8 =$	256
$0 \times 2^7 =$	0
$0 \times 2^6 =$	0
$0 \times 2^5 =$	0
$0 \times 2^4 =$	0
$1 \times 2^3 =$	8
$1 \times 2^2 =$	4
$0 \times 2^1 =$	0
$1 \times 2^0 =$	1
	6 413

En prenant successivement les multiplicateurs de la colonne de gauche (comme on le fait dans le système décimal en prenant successivement les multiplicateurs 6, 4, 1, et 3), nous écrivons le nombre dans le système binaire comme suit : 1100100001101.

A première vue, cette procédure paraît bien compliquée. Il faut 13 chiffres pour écrire le nombre 6 413, alors que le système décimal n'en a besoin que de quatre. Mais pour une machine à calculer, ce système est à peu près le plus simple qu'on puisse imaginer. Puisqu'il n'implique que deux chiffres, n'importe quelle opération peut s'effectuer en termes de oui ou non.

Il ne doit pas être sorcier de concevoir un dispositif quelconque, en s'inspirant par exemple de la présence ou de l'absence d'aiguilles dans le métier de Jacquard, pour reproduire ce oui et ce non, ou le 1 et le 2. En usant certaines combinaisons ingénieuses, on peut s'arranger pour obtenir

$0 + 0 = 0$, $0 + 1 = 1$, $0 \times 0 = 0$, $0 \times 1 = 0$ et $1 \times 1 = 1$. Une fois que de telles combinaisons sont possibles, on doit pouvoir effectuer tous les calculs arithmétiques à l'intérieur d'un dispositif du type métier de Jacquard.

Et on peut même imaginer d'autres applications. Pourquoi ne pas élargir le système et inclure les opérateurs logiques, que nous avons souvent tendance à oublier et qui font pourtant partie intégrante des mathématiques ?

En 1936, le mathématicien anglais Alan Mathison Turing a montré que n'importe quel problème peut être résolu mécaniquement si on arrive à le représenter sous la forme d'un nombre fini d'opérations réalisables par une machine.

En 1938, le mathématicien et ingénieur américain Claude Elwood Shannon a démontré dans sa thèse que le raisonnement déductif, sous la forme de l'*algèbre de Boole*, est manipulable au moyen du système binaire. L'algèbre de Boole est un système de *logique symbolique* proposé en 1854 par le mathématicien anglais George Boole dans un livre intitulé *An Investigation of the Laws of Thought* (Une investigation des lois de la pensée). Boole a observé qu'on pouvait utiliser des symboles mathématiques pour représenter les arguments du raisonnement déductif, et il a montré comment de tels symboles pouvaient être manipulés selon des règles fixes pour parvenir à formuler certaines conclusions.

Pour en donner un exemple très simple, examinons l'argument suivant : « A et B sont tous les deux vrais. » Il nous revient de déterminer logiquement si cet argument est vrai ou faux, en supposant que nous savons si A et B sont, respectivement, vrais ou faux. Pour résoudre ce problème en termes binaires, comme Shannon le suggérait, admettons que 0 soit « faux » et 1 soit « vrai ». Si A et B sont tous les deux faux, l'argument « A et B sont tous les deux vrais » est faux. En d'autres termes, 0 et 0 donnent 0. Si A est vrai et B est faux (ou vice versa), alors l'argument est faux de nouveau : c'est-à-dire que 1 et 0 (ou 0 et 1) donnent 0. Si A est vrai et B est vrai, alors l'argument « A et B sont tous les deux vrais » est vrai. De manière symbolique, 1 et 1 font 1.

Ces trois alternatives correspondent aux trois possibilités de multiplications dans le système binaire : $0 \times 0 = 0$, $1 \times 0 = 0$, et $1 \times 1 = 1$. Donc, le problème de logique posé par l'argument « A et B sont tous les deux vrais » peut être résolu grâce à la multiplication. Une machine (convenablement programmée) est ainsi capable de résoudre ce petit problème aussi facilement, et de la même façon, qu'elle effectue les opérations classiques.

Dans le cas de l'argument « Soit A soit B est vrai », le problème est résolu par l'addition au lieu de la multiplication. Si ni A ni B ne sont vrais, l'argument est alors faux. Autrement dit, $0 + 0 = 0$. Si A est vrai et B faux, ou vice versa, l'argument est vrai ; dans ces cas-là, on a $1 + 0 = 1$ et $0 + 1 = 1$. Si A et B sont tous les deux vrais, l'argument est certainement vrai, et $1+1=10$. (Le chiffre significatif dans le 10 est le 1 ; le fait qu'il soit déplacé d'un cran importe peu. Dans le système binaire, 10 signifie $[1 \times 2^1] + [0 \times 2^0]$, qui équivaut à 2 dans le système décimal.)

L'algèbre booléenne est très importante dans les techniques de communications et fait partie de ce qu'on appelle aujourd'hui la *théorie de l'information*.

L'intelligence artificielle

Le premier à saisir pleinement tout le potentiel qu'on pouvait tirer des cartes perforées du métier à tisser de Jacquard fut le mathématicien anglais Charles Babbage. En 1823, il entreprit la construction d'une machine qu'il appela « machine différentielle », puis en 1836, celle d'une « machine analytique » encore plus complexe, mais ne les finit ni l'une ni l'autre.

En théorie, ses idées étaient tout à fait valables. Il pensait pouvoir effectuer automatiquement des opérations mathématiques grâce à des cartes perforées, puis voir les résultats imprimés ou encore perforés sur des cartes vierges. Il pensait également pouvoir doter sa machine d'une mémoire en stockant des cartes préalablement perforées selon un code établi, puis en les rappelant pour des utilisations ultérieures.

La mécanique de la machine devait être constituée d'un assemblage de bielles, de cylindres, d'engrenages et de roues diverses, le tout assemblé selon une architecture interne basée sur le système décimal. Des sonnettes auraient prévenu les assistants quand il fallait introduire des cartes, et d'autres sonnettes, plus stridentes, les auraient avertis qu'une carte incorrecte avait été introduite.

Babbage, hélas, personnage excentrique et fougueux, démontait périodiquement ses machines et les reconstruisait de manière encore plus compliquée pour tenir compte de ses dernières trouvailles, et inévitablement, l'argent vint un jour à manquer.

De toute façon, les éléments mécaniques qu'il utilisait dans ses machines ne furent jamais à la hauteur des tâches qu'il leur demandait. Les machines de Babbage exigeaient une technologie plus avancée que celle de la machine de Pascal, et cette technologie n'existait tout simplement pas à son époque.

Pour toutes ces raisons, les travaux de Babbage furent interrompus et oubliés pendant un siècle. C'est seulement après avoir redécouvert tous ces principes qu'on parvint enfin à construire des machines à calculer du même type.

LES CALCULATEURS ÉLECTRONIQUES

C'est grâce aux exigences du recensement aux États-Unis qu'a pu se développer l'application pratique des cartes perforées. La constitution de ce pays impose l'organisation d'un recensement tous les dix ans, et les études statistiques menées sur sa population et son économie se sont révélées extrêmement précieuses. En effet, la population et la richesse du pays ont augmenté régulièrement tous les dix ans, et le besoin d'analyses statistiques détaillées a augmenté proportionnellement. Mais le temps passé à dépouiller toutes ces statistiques devenait de plus en plus prohibitif. Dans les années 1880, il sembla que le recensement du début de cette décade ne serait pas « bouclé » avant que celui de 1890 ne commence.

C'est alors que Herman Hollerith, un statisticien qui travaillait pour le bureau du recensement, développa un système destiné à compiler les statistiques au moyen de cartes perforées mécaniquement à des positions déterminées. Les cartes n'étaient pas elles-même conductrices, mais des impulsions électriques passaient par des contacts situés dans les trous ; de

cette façon, le comptage et les autres opérations s'effectuaient automatiquement grâce au courant électrique – un progrès considérable, et même crucial par rapport aux relais purement mécaniques de la machine de Babbage. L'électricité était à la hauteur de la tâche !

Aux États-Unis, les recensements de 1890 et 1900 firent un usage intensif de la machine électromécanique d'Hollerith. Le recensement de 1890, qui portait sur soixante-cinq millions d'habitants, mit deux ans et demi à être dépouillé, même avec les machines d'Hollerith. En 1900, il leur avait apporté d'importantes améliorations, puisqu'elles étaient désormais alimentées en cartes automatiquement ; les statistiques du recensement de 1900 sortirent en un peu plus d'un an et demi.

Hollerith fonda une compagnie qui allait devenir quelques années plus tard International Business Machines (IBM). Cette nouvelle société, ainsi que Remington Rand, fondée par l'assistant d'Hollerith, John Powers, a constamment apporté des améliorations aux systèmes de calcul électromécaniques au cours des trente années qui ont suivi.

Et il était grand temps.

L'économie mondiale, avec son industrialisation galopante, allait vers une complexité toujours croissante ; de plus en plus, sa gestion exigeait une connaissance aussi complète que possible des statistiques la concernant ; il fallait des chiffres, il fallait des informations. Le monde devenait une *société d'informations* qui risquait de s'écrouler si l'homme ne parvenait pas rapidement à obtenir, comprendre, et maîtriser ces informations.

C'est sous l'effet de cette pression impitoyable, qui *rendait inéluctable* la manipulation de quantités croissantes d'informations, que la société a été amenée à inventer au XXe siècle des machines à calculer de plus en plus sophistiquées et variées.

La vitesse d'exécution des machines électromécaniques n'a cessé d'augmenter, et leur utilisation s'est étendue jusqu'à la fin de la Deuxième Guerre mondiale ; mais leur vitesse et leur fiabilité étaient limitées par l'utilisation de pièces mobiles telles que les contacteurs et les relais électromagnétiques qui actionnaient les roues des compteurs.

En 1925, l'ingénieur américain Vannevar Bush et ses collègues ont construit pour la première fois une machine capable de résoudre des équations différentielles. Elle faisait tout ce que Babbage espérait de sa propre machine, et elle représente véritablement le premier instrument que l'on qualifierait aujourd'hui d'*ordinateur*. Mais elle était encore composée d'éléments électromécaniques.

Toujours de conception électromécanique, mais encore plus impressionnante, fut la machine mise au point en 1937 par Howard Aiken, de Harvard, en collaboration avec IBM. La machine, le « calculateur automatique contrôlé par séquence d'IBM », connue à Harvard sous le nom de Mark I, fut terminée en 1944 ; elle était destinée à des applications scientifiques, et effectuait des opérations mathématiques avec une précision de vingt-trois décimales. Grâce à elle, on pouvait multiplier deux nombres de onze chiffres et obtenir le résultat en trois secondes. Elle aussi était électromagnétique ; et comme elle travaillait principalement sur des nombres, il s'agit là du premier exemple de *calculateur digital*. (L'appareil de Bush résolvait les problèmes en convertissant les nombres en longueur, comme le fait une règle à calcul ; et comme il se basait sur des *quantités analogues*, et non pas vraiment des nombres, c'était un *calculateur analogique*.)

Mais pour devenir fiables, ces machines ont dû attendre l'avènement de l'électronique. Les interruptions mécaniques et électriques inhérentes à ce type de relais, bien qu'offrant des performances infiniment supérieures à celles de roues et d'engrenages, étaient tout de même sources de lenteur et de lourdeur, sans parler des nombreuses pannes qu'elles pouvaient occasionner. Dans les relais électroniques, les tubes cathodiques par exemple, il devenait possible de maîtriser le flot des électrons beaucoup plus délicatement, avec plus de précision et infiniment plus rapidement ; ce fut l'étape suivante.

Le premier gros calculateur électronique, qui contenait 19 000 tubes cathodiques, a été construit à l'université de Pennsylvannie par John Presper Eckert et John William Mauchly pendant la Deuxième Guerre mondiale. On l'a baptisé ENIAC, un acronyme tiré de l'expression anglaise *E*lectronic *N*umerical *I*ntegrator *a*nd *C*omputer. ENIAC a pris sa retraite en 1955 et on l'a démonté deux ans plus tard : il était devenu un pauvre vieux radoteur irrécupérable et démodé à l'âge de douze ans ; mais il a laissé derrière lui une descendance incroyablement prolifique et sophistiquée. Alors qu'ENIAC pesait trente tonnes et prenait une place considérable (quelque cent cinquante mètres cubes au sol), on construit trente ans plus tard des ordinateurs aux performances comparables, mais dont les dimensions sont celles d'un réfrigérateur ; l'ordinateur d'aujourd'hui utilise des composants miniaturisés plus rapides et plus fiables que les tubes cathodiques d'antan.

Les progrès en la matière furent si étonnants que plusieurs dizaines de petits calculateurs électroniques existaient déjà en 1948 ; cinq ans plus tard, il y en avait deux mille ; en 1961, leur nombre avoisinait les dix mille. On en comptait cent mille en 1970, et il ne s'agissait là que d'un début.

La raison de ce développement foudroyant tient au fait que, si l'électronique offrait effectivement une planche de salut, le tube cathodique, lui, n'était plus à la hauteur des tâches demandées. Il était volumineux, fragile, gourmant en énergie. En 1948 s'ouvrait l'ère du transistor (voir chapitre 9) ; avec elle, l'électronique devenait compacte, solide, et faisait faire des économies d'énergie.

Les ordinateurs devenaient plus petits et moins chers alors que leur capacité-mémoire et leur souplesse d'utilisation augmentaient considérablement. Au cours des vingt-cinq ans qui ont suivi l'invention du transistor, une succession impressionnante de nouvelles techniques a permis de bourrer de plus en plus d'informations et de mémoire dans des composants de plus en plus minuscules. En 1970 est apparu le *microchip* – un petit morceau de silicium sur lequel, sous l'œil d'un microscope, on peut imprimer des milliers de circuits.

Le résultat ne s'est pas fait attendre : tout le monde aujourd'hui, sans qu'il soit besoin d'être millionnaire, peut s'offrir un ordinateur. Il est possible que les années quatre-vingts deviennent pour l'*ordinateur familial* ce que les années cinquante ont été pour les postes de télévision.

Pour le grand public, les calculateurs qui se sont répandus après la Deuxième Guerre mondiale ressemblaient déjà à des « machines pensantes », si bien qu'à la fois des chercheurs et des non-scienfitiques se sont mis à réfléchir aux possibilités, et aux conséquences, de l'*intelligence artificielle*, une expression utilisée pour la première fois en 1956 par John McCarthy, ingénieur en électronique au MIT (Institut de technologie du Massachusetts).

En moins de quarante ans, les ordinateurs sont devenus des géants sans lesquels notre civilisation s'écroulerait : l'exploration spatiale serait impossible sans les ordinateurs ; la navette ne décollerait pas sans eux. Notre machine de guerre serait telle qu'elle était pendant la Deuxième Guerre mondiale. Aucune industrie, quelle que soit sa taille, pas même le plus petit bureau, ne pourrait continuer de fonctionner comme aujourd'hui. Le gouvernement serait encore plus impotent qu'il ne l'est habituellement.

De nouvelles applications se développent tous les jours. Non seulement les ordinateurs calculent, dessinent des graphiques, accumulent et stockent des données, etc., mais ils sont aussi capables d'accomplir des tâches triviales. On les programme pour jouer aux échecs (presque) aussi bien que des maîtres, ou pour d'autres jeux de toutes sortes qui ont créé un tel engouement dans la jeunesse actuelle que ce sont littéralement des milliards de dollars qu'on leur consacre chaque année. Les informaticiens travaillent actuellement sur des programmes qui permettront la traduction automatique d'une langue vers une autre, et donneront aux ordinateurs la faculté de lire, d'entendre et de parler.

LES ROBOTS

La question suivante se pose inévitablement : les ordinateurs peuvent-ils, en dernière analyse, tout faire ? Notre imagination n'est-elle pas, finalement, leur seule limite ? Ne peut-on pas imaginer, par exemple, un ordinateur spécialement conçu, habillé d'une enveloppe ressemblant au corps humain ? Nous aurions ainsi un véritable automate – pas un de ces jouets du XVIIe siècle, mais un être humain artificiel doué de presque toutes nos facultés.

Une telle éventualité a été envisagée très sérieusement par les auteurs de science-fiction avant même que soient fabriqués les premiers ordinateurs modernes. En 1920, l'auteur dramatique tchèque Karel Capek a écrit *R. U. R.*, une pièce dans laquelle il raconte l'histoire d'un Anglais, Rossum, qui fabrique des automates à la chaîne, dans le but de leur faire accomplir les tâches les plus ingrates et de rendre ainsi la vie des hommes plus facile. A la fin, ils se révoltent, exterminent l'humanité tout entière, et fondent une nouvelle race d'êtres intelligents.

Rossum vient d'un mot tchèque, *rozum*, qui signifie « raison » ; et *R. U. R.* sont les initiales de « Robots Universels de Rossum », une expression dans laquelle *robot* est un mot tchèque qui veut dire « travailleur », mais avec une forte connotation de servitude ; on peut le traduire par « serf » ou « esclave ». La pièce remporta un tel succès que le vieux terme automate tomba en désuétude. Le mot *robot* l'a remplacé dans toutes les langues, si bien qu'on imagine aujourd'hui un robot sous la forme d'un engin artificiel (souvent d'apparence vaguement humaine) qui, grosso modo, se comporte comme le feraient des êtres humains.

Dans l'ensemble, les auteurs de science-fiction ont rarement utilisé les robots sous cette forme, mais plutôt sous les traits de héros ou de traîtres, pour mieux souligner les particularités de la condition humaine.

C'est alors qu'en 1939, Isaac Asimov (eh oui, l'auteur de cet ouvrage !), dix-neuf ans à peine révolus, fatigué par des robots qui étaient soit d'une méchanceté incroyable, soit d'une noblesse non moins incroyable, décida

qu'il était grand temps d'écrire des histoires de science-fiction sur des robots qui ne seraient tout simplement que des machines, fabriqués comme le sont toutes les machines, c'est-à-dire en prenant certaines précautions que dicte le bon sens, tout simplement. Il publia bon nombre de ces histoires sous forme de nouvelles dans les années quarante ; et en 1950 neuf d'entre elles sont parues dans un recueil sous le titre *Les Robots*.

Les précautions d'Asimov ont donné lieu aux « Trois Lois de la Robotique ». Elles apparaissent pour la première fois dans une nouvelle publiée en mars 1942, la première fois au monde que le terme *robotique* est utilisé, terme maintenant universellement accepté qui définit la science et la technologie du développement, de la construction, de la maintenance et de l'utilisation des robots.

Les trois règles sont les suivantes :

1 Un robot ne doit pas faire de mal à un être humain ni, par son inaction, laisser faire du mal à un être humain.

2 Un robot doit obéir aux ordres que lui donne un être humain, à moins que ces ordres ne soient en conflit avec la Première Loi.

3 Un robot doit protéger son existence aussi longtemps que cette protection ne constitue pas un conflit avec la Première ou la Seconde Loi.

C'était là, bien évidemment, l'imagination d'Asimov qui parlait, et qui ne pouvait servir, au mieux, que d'inspiration pour les travaux sérieux des scientifiques dans ce domaine.

Les pressions de la Deuxième Guerre mondiale ont joué un rôle de premier plan. Les applications de l'électronique ont permis la fabrication d'armes tellement sophistiquées qu'elles dépassent, et de loin, les capacités de n'importe quel organisme vivant. La fusée V1 allemande était essentiellement un *servomécanisme*, qui a ouvert le chemin non seulement aux missiles guidés, mais aussi à toutes sortes de véhicules à fonctionnement automatique ou télécommandés, du métro à la navette spaciale. A cause de l'intérêt très vif porté à ce genre d'engins par les autorités militaires et grâce aux fonds très importants dont celles-ci disposent, les servomécanismes ont atteint leur plus haut degré de sophistication dans le domaine des systèmes de guidage et de mise-à-feu pour les missiles et les fusées. Ces systèmes détectent une cible volant à grande vitesse à plusieurs centaines de kilomètres, calculent instantanément sa trajectoire (en tenant compte de la vitesse de la cible, du vent, de la température des différentes couches de l'atmosphère, et de nombreuses autres conditions de tous ordres), et la détruisent avec une précision diabolique, le tout sans la moindre assistance humaine.

L'automatisation a trouvé en la personne du mathématicien Norbert Wiener un ardent défenseur et un théoricien de tout premier plan. Ce dernier a pris une part prépondérante dans l'élaboration des systèmes de guidage. Dans les années quarante, lui et son équipe du MIT ont découvert certaines des relations mathématiques fondamentales impliquées dans le feedback. Ils ont appelé ce domaine la *cybernétique*, du mot grec qui signifie « timonier », ce qui paraît très approprié puisque la première application d'un servomécanisme était liée à un timonier. (Le terme connote également le gouverneur centrifuge de Watt, et le mot *gouverneur* vient du latin « timonier ».)

Le livre qu'ils ont publié était le premier entièrement consacré à la théorie de l'architecture interne des ordinateurs, et les principes de la cybernétique ont ouvert la voie, peut-être pas aux robots, mais à des

systèmes qui utilisent ces règles pour imiter le comportement de certains animaux.

Le neurologue anglais William Grey Walter, par exemple, a construit un appareil, dans les années cinquante, qui explore et réagit à son environnement. L'objet ressemble à une tortue, et il l'a appelé *testudo* (du latin « tortue ») ; il comporte une cellule photo-électrique à la place de l'œil, des capteurs pour le toucher, et deux moteurs – un pour se déplacer vers l'avant et vers l'arrière, l'autre pour se retourner. Dans le noir, il bouge lentement en faisant de grands cercles. Lorsqu'il heurte un obstacle, il recule, tourne légèrement, puis avance de nouveau ; cette même manœuvre se poursuit jusqu'à ce qu'il ait réussi à contourner l'obstacle. Lorsque son œil photo-électrique perçoit une lumière, le moteur qui lui permet de tourner s'arrête de fonctionner, et le testudo se dirige droit vers la lumière. Mais son phototropisme est contrôlé ; lorsqu'il s'approche de la lumière, l'intensité plus forte lui fait faire marche arrière, ce qui lui évite de commettre l'erreur du papillon de nuit. Mais quand ses piles commencent à se décharger, le testudo qui a alors « faim » s'approche de la lumière où se trouve une prise à laquelle il peut se brancher et refaire son plein d'énergie. Une fois rechargé, le testudo est de nouveau assez sensible pour s'éloigner de la zone trop éclairée.

Mais il ne faut surtout pas sous-estimer l'influence que peut avoir l'inspiration. Au début des années cinquante, un jeune étudiant de l'université Columbia, Joseph F. Engelberger, lut les *Les robots*, et attrapa le virus des robots.

En 1956, il rencontra George C. Devol Jr., qui, deux ans auparavant, avait obtenu le premier brevet industriel pour la construction d'un robot. Il avait baptisé son système de contrôle et de mémoire *universal automation* (automatisation universelle), ou *unimation*, en abrégé.

Ensemble, Engelberger et Devol fondèrent la compagnie Unimation, et Devol développa entre trente et quarante brevets différents dans ce domaine.

Aucun de ces brevets n'avait vraiment d'application pratique, car les robots ne pouvaient accomplir aucune tâche sans être préalablement informatisés, et les ordinateurs étaient encore trop volumineux et trop coûteux à cette époque. Le robot n'était pas encore une option économiquement compétitive. Ce n'est que grâce à l'avènement de la puce électronique que les robots d'Unimation ont pu trouver leur place sur le marché. Unimation est bientôt devenue l'entreprise de robotique la plus importante et la plus prospère au monde.

L'ère du *robot industriel* était née. Le robot industriel ne ressemble pas au robot classique ; il n'a rien de particulièrement humanoïde. C'est surtout un bras informatisé, qui peut effectuer des travaux simples avec une grande précision et qui jouit, grâce à l'informatique, d'une certaine flexibilité.

Aujourd'hui, on utilise surtout les robots industriels sur les chaînes de montage (particulièrement au Japon, dans l'industrie automobile). Pour la première fois, nous possédons des machines suffisamment complexes et « talentueuses » pour accomplir des tâches qui jusqu'à maintenant revenaient seulement à l'homme grâce à son intelligence – mais nécessitant en fait si peu d'intelligence que le cerveau, forcé d'accomplir un travail généralement répétitif et abêtissant, non seulement gâchait son potentiel, mais risquait probablement par là-même de se détériorer.

L'utilité de telles machines est claire : elles permettent d'exécuter des travaux trop routiniers ou trop pénibles pour l'homme, ce qui lui laisse le temps pour faire des choses plus créatives.

On s'aperçoit hélas que leur utilisation provoque à court terme certains effets néfastes : elles remplacent en effet certains postes de travail. Il va nous falloir affronter une difficile période de transition au cours de laquelle la société devra prendre en charge les nouveaux chômeurs, et les former à d'autres métiers ; ou lorsque cela sera impossible, leur trouver des occupations utiles ; ou enfin, si rien de tout cela n'est faisable, subvenir à leurs besoins.

Il faut espérer qu'avec le temps, une nouvelle génération rompue aux techniques de l'informatique et de la robotique prendra le relais, et que la situation s'améliorera.

Mais les progrès de la technologie sont inexorables. On pousse de plus en plus à la fabrication de robots toujours plus sophistiqués, qui « voient », « parlent », « entendent ». Qui plus est, il existe maintenant des *robots domestiques* – des robots à l'apparence plus humanoïde qui accomplissent des tâches ménagères habituellement confiées aux serviteurs humains. (Joseph Engelberger a mis au point un engin dont il espère essayer le prototype chez lui sous peu : il prend les manteaux, sert à boire, et accomplit d'autres petites tâches simples. Il l'appelle Isaac.)

Nous devons nous demander si les ordinateurs et les robots ne copieront pas un jour *n'importe* quelle activité humaine ; s'ils ne nous rendront pas obsolètes ; si l'intelligence artificielle, que nous aurons ainsi créée, n'est pas destinée à nous remplacer sur cette planète.

Nous pourrions réagir de manière fataliste. Si c'est inévitable, eh bien c'est inévitable. En plus, qu'avons-nous vraiment accompli de bien sur cette Terre ? Nous sommes peut-être en train de nous détruire (et toute forme de vie avec nous), de toute façon ; le problème n'est pas de savoir si nous serons remplacés par les ordinateurs, le problème est de savoir quand cela arrivera.

Nous pourrions réagir triomphalement. Quel superbe exploit que de créer un objet qui surpasse son créateur ! Comment consommer la victoire de l'intelligence plus glorieusement qu'en léguant notre héritage à une intelligence supérieure – qui nous devrait la vie ?

Mais soyons objectifs. Courons-nous réellement le danger d'être un jour remplacés par des robots ?

Tout d'abord, nous devons nous poser la question de savoir si l'intelligence est une variable à une dimension, ou s'il n'y aurait pas plusieurs sortes d'intelligence qualitativement différentes, et même un très grand nombre d'intelligences, toutes différentes. Bien que les dauphins possèdent une intelligence tout à fait semblable à la nôtre, par exemple, il semble qu'elle soit d'une nature si différente que nous sommes encore actuellement incapables de communiquer avec eux. Il se peut, en dernière analyse, que les ordinateurs soient également différents de nous qualitativement. Si tel était le cas, cela ne devrait certainement pas nous surprendre.

Après tout, le cerveau humain, fabriqué à partir d'acides nucléiques et de protéines sur fond d'H_2O, est le produit d'une évolution biologique qui a commencé voilà trois milliards et demi d'années, et qui est basée sur les effets aléatoires des mutations, de la sélection naturelle, ainsi que d'autres phénomènes, le tout dans un seul but : survivre.

L'ordinateur, de son côté, fabriqué à partir de composants électroniques et de microchips sur fond de semiconducteurs, est le produit de quarante

ans d'invention humaine, basée sur la prudence et l'ingéniosité de l'homme, le tout dans un seul but : servir ce dernier.

Si deux intelligences ont des structures, des histoires, des développements et des buts si différents, cela ne surprendra personne que ces deux intelligences soient également de natures très différentes.

Depuis leurs débuts, les ordinateurs sont capables de résoudre des problèmes très complexes d'arithmétique impliquant des opérations sur des nombres, beaucoup plus rapidement que n'importe quel être humain ne pourrait le faire, et avec bien moins de risque d'erreur. Si les compétences en arithmétique sont la mesure de l'intelligence, les ordinateurs ont toujours été plus intelligents que nous.

Mais il est tout à fait possible que les compétences en arithmétique, ou dans d'autres domaines du même genre, ne soient pas particulièrement ce pour quoi notre cerveau est le plus doué – nous sommes loin de briller dans ces disciplines que nous n'avons pas l'habitude de pratiquer.

Il se peut que la mesure de l'intelligence implique des qualités aussi subtiles que la perspicacité, l'intuition, la fantaisie, l'imagination, la création – la capacité à appréhender un problème dans son ensemble et à deviner la réponse grâce au contexte qui l'entoure, au « feeling » de la situation. S'il en est ainsi, les êtres humains sont alors très intelligents, et les ordinateurs très stupides. Et nous ne voyons pas facilement aujourd'hui comment remédier à ce défaut des ordinateurs, car nous ignorons comment les programmer pour qu'ils deviennent intuitifs ou créatifs, et ceci pour la simple et bonne raison que nous ne savons pas nous-mêmes comment diable ces qualités s'exercent chez nous.

Nous sera-t-il un jour possible de programmer des ordinateurs à de telles fins ?

La chose est concevable ; mais il se peut aussi que nous choisissions de ne pas le faire de crainte d'être un jour supplantés. De plus, quelle serait l'utilité de copier l'intelligence humaine – de construire un ordinateur qui aurait un mince vernis d'humanité – quand il nous est si facile de la fabriquer au moyen de processus biologiques ordinaires ? Ce serait un peu comme si on dressait des petits enfants dès la naissance à imiter les prouesses mathématiques des ordinateurs. Pourquoi faire cela, alors qu'une calculatrice bon marché sait très bien le faire ?

Nous ne pourrions que tirer un grand bénéfice du développement simultané de deux intelligences aux spécialités différentes, permettant ainsi à plus de fonctions de se dérouler avec un maximum d'efficacité. On pourrait même imaginer plusieurs types d'ordinateurs caractérisés par des intelligences différentes. Et en ayant recours aux méthodes du génie génétique (et bien sûr avec l'aide des ordinateurs), il serait possible de créer différents types de cerveaux humains caractérisés par différents types d'intelligences humaines.

Avec toutes ces différentes espèces d'intelligences, la possibilité d'une relation symbiotique existe, dans laquelle toutes coopéreront pour apprendre à mieux connaître les lois de la nature et à mieux les utiliser. Il est probable qu'une telle coopération donnera de meilleurs résultats que n'importe quelle variété unique d'intelligence.

Vu sous cet angle, le robot/ordinateur ne nous remplacera pas, mais deviendra notre ami et notre allié tout au long de la route qui nous conduit vers un futur glorieux – si nous ne nous détruisons pas avant même que le voyage ne commence.

Appendice

Le rôle des mathématiques

La gravitation

Comme nous l'avons vu au chapitre 1, c'est Galilée qui a donné le départ à la science moderne, en introduisant l'idée de remonter des expériences et des observations aux principes fondamentaux, par le raisonnement. Ce faisant, il a aussi introduit la mesure précise des phénomènes naturels et a abandonné l'habitude de se contenter de les décrire. En somme, il a remplacé la description qualitative de l'Univers, telle que la donnaient les anciens Grecs, par une description quantitative.

Bien que la science dépende des relations et des manipulations mathématiques au point qu'elle ne pourrait pas, sans elles, exister au sens galiléen, j'ai écrit ce livre d'une manière non mathématique, et cela délibérément. Après tout, les mathématiques sont un outil spécialisé. Pour exposer les progrès scientifiques en termes mathématiques, il aurait fallu beaucoup de place, et aussi – de la part du lecteur – une connaissance étendue des mathématiques. Mais dans cet appendice, j'aimerais présenter un exemple ou deux d'application fructueuse à la science de techniques mathématiques simples. Et comment commencer, sinon avec Galilée lui-même ?

LA PREMIÈRE LOI DU MOUVEMENT

Galilée (comme Léonard de Vinci presque cent ans plus tôt) pressentait que les objets vont de plus en plus vite au cours d'une chute. Il a cherché à mesurer exactement de combien, et de quelle manière, leur vitesse augmente.

Avec les instruments dont il disposait en 1600, cette tâche n'était pas du tout facile. Pour mesurer une vitesse, il faut mesurer le temps. Nous parlons de vitesses de 72 kilomètres *à l'heure* ou de 20 mètres *par seconde*. Mais à l'époque de Galilée, les horloges étaient tout juste capables de sonner grosso modo toutes les heures...

Galilée s'est rabattu sur une horloge à eau rudimentaire. Il laissait l'eau s'échapper lentement par un petit trou et supposait, d'une façon un peu optimiste, qu'elle s'écoulait d'une façon uniforme. Cette eau qui s'échappait, il la recueillait dans un bol et la pesait pour savoir combien de temps avait duré tel ou tel phénomène. (Il lui est aussi arrivé d'utiliser son pouls à cet effet.)

Mais les objets tombent si vite que la quantité d'eau recueillie pendant leur chute était trop faible pour pouvoir être pesée de façon assez précise. Galilée a donc *dilué* la pesanteur en faisant rouler une bille dans une rainure le long d'un plan incliné. Plus le plan était proche de l'horizontale, plus la bille se déplaçait lentement. Ainsi, Galilée pouvait étudier des *mouvements* aussi *ralentis* qu'il le souhaitait.

Il a constaté qu'une bille roulait à vitesse constante sur un plan horizontal (ce qui suppose l'absence de frottements, condition assez facilement acceptable compte tenu de la faible précision des mesures). Or, un corps qui se déplace horizontalement se déplace perpendiculairement à la force exercée par la gravitation. Dans ces conditions, celle-ci ne change pas la vitesse. Une balle immobile sur un plan horizontal reste immobile, comme chacun sait. Une balle mise en mouvement sur un plan horizontal garde une vitesse constante, comme Galilée l'a observé.

Mathématiquement, on peut donc affirmer qu'*en l'absence de force extérieure* la vitesse v d'un corps garde une valeur constante k soit :

$$v = k$$

Si k est égal à un nombre différent de zéro, la bille se déplace à vitesse constante. Si k est égal à zéro, la bille est immobile. Ainsi, l'immobilité est un « cas particulier » du déplacement à vitesse constante.

Presque un siècle plus tard, quand Newton systématisa les découvertes faites par Galilée au sujet de la chute des corps, cette découverte est devenue la « première loi du mouvement » (ou *principe de l'inertie*). On peut énoncer cette loi de la manière suivante : tout corps reste immobile ou animé d'un mouvement rectiligne uniforme en l'absence d'une force extérieure.

Mais quand la bille descend le long d'un plan incliné, elle est constamment soumise à l'attraction gravitationnelle. Dans ces conditions, comme l'a montré Galilée, sa vitesse n'est pas constante, mais augmente avec le temps. D'après les mesures de Galilée, cette vitesse augmente proportionnellement au temps t qui s'écoule.

En d'autres termes, quand un corps est soumis à une force extérieure constante, sa vitesse, s'il part du repos, peut s'exprimer ainsi :

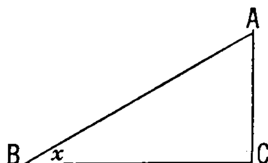
$$v = kt$$

Quelle est la valeur de k ?

Comme l'expérience l'a facilement montré, cette valeur dépend de la pente du plan incliné. Plus le plan est incliné, plus la bille gagne rapidement de la vitesse, et plus la valeur de k est grande. Le maximum est atteint quand le plan est vertical, autrement dit quand la bille tombe en chute libre, sous l'effet d'une attraction gravitationnelle non diluée. Dans ce cas, on utilise la lettre g (pour « gravitation »), et la vitesse d'une bille en chute libre, un temps t après son départ (départ arrêté), s'écrit :

$$v = gt$$

Regardons de plus près le plan incliné. Sur le schéma ci-dessous,



la longueur du plan incliné est AB, tandis que sa hauteur à l'extrémité la plus haute est AC. Le rapport AC/AB est le sinus de l'angle x qui s'écrit généralement en abrégé « sin x ».

La valeur de ce rapport, c'est-à-dire de sin x, peut s'obtenir approximativement en construisant des triangles d'angles divers, et en mesurant effectivement la hauteur et la longueur en question. Ou bien on peut la calculer mathématiquement, avec la précision que l'on veut, et dresser un tableau des résultats. En utilisant ce tableau, on peut trouver que, par exemple, sin 10^0 vaut à peu près 0,17365, que sin 45^0 vaut à peu près 0,70711, et ainsi de suite.

Il y a deux cas particuliers importants. Supposons que le plan soit exactement horizontal. L'angle x est alors nul, la « hauteur » du plan incliné est nulle aussi, donc le rapport de cette « hauteur » et de la longueur du plan est également nul. Autrement dit, $\sin 0^0 = 0$. Quand le plan « incliné » est exactement vertical, il forme avec le sol un angle droit, soit 90^0. Sa hauteur est alors exactement égale à sa longueur, et leur rapport vaut 1. Par conséquent, $\sin 90^0 = 1$.

Revenons à l'équation qui dit que la vitesse d'une bille descendant le long d'un plan incliné est proportionnelle au temps :

$$v = kt$$

On peut montrer expérimentalement que la valeur de k change avec le sinus de l'angle, de telle manière que :

$$k = k' \sin x$$

(k' étant une constante différente de k).

(En fait, le rôle du sinus de l'inclinaison du plan avait été établi quelque temps avant l'époque de Galilée par Simon Stevinus, qui avait également réalisé l'expérience célèbre consistant à laisser tomber de haut différentes masses, expérience traditionnellement attribuée à Galilée. Mais si Galilée n'a pas vraiment été le premier à faire des expériences et des mesures, il a été le premier à imposer au monde scientifique, d'une manière définitive, la nécessité des expériences et des mesures, et ce fut là un véritable succès.)

Dans le cas d'un plan incliné vertical, sin x devient sin 90^{0}, qui vaut 1, c'est-à-dire qu'en chute libre :

$$k = k'$$

Donc k' est la valeur de k en chute libre, sous l'effet de l'attraction gravitationnelle tout entière, c'est-à-dire la constante que nous avons décidé d'appeler g. Nous pouvons donc remplacer k' par g et pour n'importe quel plan incliné :

$$k = g \sin x$$

Par conséquent, la vitesse d'un corps roulant sur un plan incliné est :

$$v = (g \sin x)\, t$$

Sur un plan horizontal, avec $\sin x = \sin 0^{0} = 0$, l'équation donnant la vitesse devient :

$$v = 0$$

C'est une autre façon de dire qu'une bille abandonnée sans vitesse sur un plan horizontal restera immobile aussi longtemps qu'on voudra. Un objet au repos tend à rester au repos, et ainsi de suite. C'est une partie de la première loi du mouvement, et elle se déduit de l'équation de la vitesse sur un plan incliné.

Supposons maintenant qu'au lieu de partir du repos, la bille ait déjà une certaine vitesse quand elle commence à tomber.

Supposons, par exemple, qu'une bille se déplace à la vitesse horizontale constante de 2 m/s, quand elle se trouve soudain en haut d'un plan incliné, et se met à descendre.

L'expérience montre que sa vitesse vaut à chaque instant 2 m/s de plus que si elle avait démarré à partir de l'immobilité. Autrement dit, l'équation du mouvement d'une bille le long d'un plan incliné peut s'écrire, plus complètement :

$$v = (g \sin x)\, t + V$$

V étant la vitesse initiale. Si l'objet part du repos, V est nul, et nous retrouvons l'équation sous la forme :

$$v = (g \sin x)\, t$$

Si maintenant nous considérons un objet, animé d'une certaine vitesse initiale, et se déplaçant sur un plan horizontal, pour lequel $\sin x = 0$, l'équation devient :

$$v = V$$

Ainsi, la vitesse de cet objet reste égale à sa vitesse initiale, quel que soit le temps qui s'écoule. C'est la première loi du mouvement, déduite encore une fois du déplacement sur un plan incliné.

La rapidité avec laquelle la vitesse change s'appelle *l'accélération*. Si, par exemple, la vitesse (en mètres par seconde) d'une bille descendant un plan incliné vaut 2, 4, 6, 8, au bout de 1, 2, 3, 4 secondes, alors l'accélération vaut deux m/s.

En chute libre, l'équation $v = gt$ montre que chaque seconde de chute augmente la vitesse de g m/s. Par conséquent, g est l'« accélération de la pesanteur ».

La valeur de g peut être obtenue par des expériences avec le plan incliné. En effet, l'équation obtenue pour celui-ci peut s'écrire :

$$g = v/ (t \sin x)$$

On peut mesurer v, t et x et par conséquent calculer g. On trouve que cette accélération vaut environ 10 m/s/s à la surface de la Terre. En chute libre, sous l'effet de la pesanteur normale à la surface de la Terre, la vitesse de chute est liée au temps par la relation :

$$v = 10\, t$$

Voilà la solution du problème que s'était posé Galilée, à savoir déterminer la vitesse de chute d'un corps et la variation de cette vitesse en fonction du temps.

On peut ensuite se demander de quelle hauteur un corps tombe en un temps donné. A partir de l'équation donnant la vitesse en fonction du temps, on peut trouver la distance en fonction du temps, par une technique de calcul qui s'appelle *intégration.* Mais ce n'est pas indispensable, car on peut trouver cette relation par l'expérience, et c'est ce qu'a fait Galilée.

Il a constaté que la distance parcourue par une bille le long d'un plan incliné était proportionnelle au carré du temps écoulé depuis le départ. Si ce temps double, la distance parcourue est multipliée par quatre, si le temps triple elle est multipliée par neuf, etc.

Pour un corps en chute libre, la distance parcourue est liée au temps écoulé par la relation :

$$d = \frac{1}{2} g t^2$$

soit (puisque $g = 10$m/s/s) :

$$d = 5\, t^2$$

Supposons maintenant qu'au lieu de tomber à partir du repos, un objet soit lancé horizontalement d'une certaine hauteur. Son mouvement va être une composition de deux mouvements – l'un horizontal et l'autre vertical.

Le mouvement horizontal ne correspond à aucune force, à part l'impulsion de départ (si on néglige le vent, la résistance de l'air, etc.), et par conséquent, suivant la première loi du mouvement, sa vitesse est constante, et la distance horizontale parcourue est proportionnelle au temps. Mais le mouvement vertical, lui, donne, comme nous venons de le voir, une distance verticale proportionnelle au carré du temps écoulé. Avant Galilée, on pensait vaguement qu'un projectile se déplaçait en ligne droite jusqu'à ce que l'impulsion de départ soit, d'une manière ou d'une autre, épuisée, après quoi l'objet tombait verticalement. Galilée a réalisé le progrès considérable consistant à *combiner* les deux mouvements.

La composition de ces deux mouvements (un mouvement horizontal proportionnel au temps et un mouvement vertical proportionnel au carré du temps) donne une courbe appelée *parabole*. Si on lance l'objet non plus horizontalement, mais obliquement, vers le haut ou vers le bas, il décrit encore une parabole.

De telles trajectoires, évidemment, s'appliquent au mouvement des projectiles, par exemple des boulets de canon. L'analyse mathématique des trajectoires, à partir des études de Galilée, a permis de calculer l'endroit

où tomberait un boulet lancé avec une certaine vitesse initiale sous un certain angle. Les gens lançaient des objets pour le plaisir, pour la chasse et pour la guerre, depuis des milliers et des milliers d'années, mais c'est grâce à Galilée, à ses mesures et à ses expériences, qu'est née la *balistique*. Ainsi, le premier succès de la science expérimentale moderne s'est immédiatement traduit par des applications militaires.

Mais il a eu aussi des conséquences importantes sur le plan théorique. L'analyse mathématique de la composition des mouvements a permis de réfuter plusieurs objections à la théorie de Copernic. En particulier, un objet lancé verticalement en l'air à partir d'une Terre en mouvement ne retombera pas en arrière, puisqu'il a en réalité deux mouvements : celui que déclenche le lancement, et celui qu'il partage avec la surface de la Terre. De même, il devient acceptable que la Terre puisse avoir deux mouvements à la fois, et tourner à la fois sur elle-même et autour du Soleil – chose impensable selon certains des adversaires des idées de Copernic.

LA DEUXIÈME ET LA TROISIÈME LOIS

Isaac Newton a étendu aux mouvements célestes les idées de Galilée sur le mouvement, et montré que le même ensemble de lois s'appliquait au ciel et à la Terre.

Il a commencé par se dire que la Lune pouvait tomber indéfiniment vers la Terre sous l'effet de l'attraction gravitationnelle de celle-ci, sans l'atteindre jamais, grâce à la composante horizontale de sa vitesse. Comme nous l'avons vu, un projectile lancé horizontalement suit une parabole incurvée vers le bas, jusqu'à ce qu'il rencontre la surface de la Terre. Mais la surface de la Terre, elle aussi, s'incurve vers le bas, puisqu'elle est sphérique. Un projectile animé d'une vitesse initiale horizontale suffisamment grande pourrait avoir une trajectoire qui s'incurve vers le bas de la même façon que la surface de la Terre, et donc en faire éternellement le tour.

Le mouvement de la Lune autour de la Terre peut se décomposer en un mouvement horizontal et un mouvement vertical. Celui-ci est tel que la Lune, en une seconde, tombe environ de 1,2 millimètre vers la Terre (pendant qu'elle parcourt à peu près un kilomètre horizontalement). La composition des deux mouvements correspond à la courbure de la Terre.

La question est de savoir si cette chute de 1,4 millimètre par seconde de la Lune est causée par la même attraction gravitationnelle qui fait parcourir, à une pomme tombant d'un arbre, 5 mètres pendant la première seconde de sa chute.

Newton s'est représenté l'attraction de la Terre s'étendant dans toutes les directions comme une vaste sphère en expansion. La surface A d'une sphère est proportionnelle au carré de son rayon r :

$$A = 4\pi r^2$$

Il s'est donc dit que la force de gravitation, se répartissant sur la surface de cette sphère, devait diminuer en fonction du carré de son rayon. C'est ce qui se produit pour l'intensité de la lumière et du son, qui est inversement proportionnelle au carré de la distance de la source. Alors, pourquoi pas pour la gravitation ?

La distance du centre de la Terre à la pomme est environ de 6 400 kilomètres. Celle qui sépare le centre de la Terre de celui de la Lune est

environ 60 fois plus grande. Par conséquent, la force de la gravitation terrestre à la distance de la Lune doit être 60^2, soit 3 600 fois plus faible qu'à la distance de la pomme. Si on divise 5 mètres par 3 600, on trouve à peu près 1,4 millimètre. Il semblait donc clair à Newton que le mouvement de la Lune est commandé par l'attraction terrestre.

Ensuite, Newton s'est mis à considérer la masse, en relation avec l'attraction gravitationnelle. D'habitude, nous confondons volontiers masse et poids. Mais le poids résulte de l'attraction terrestre. Si celle-ci n'existait pas, les objets seraient sans poids, mais ils contiendraient toujours autant de matière. Par conséquent, la masse est distincte du poids, et doit pouvoir se mesurer par des moyens qui ne mettent pas le poids en jeu.

Supposons que nous essayions de tirer un objet le long d'une surface horizontale parfaitement lisse (sans frottements). Dans un tel mouvement horizontal, la gravitation n'interviendrait pas. Mais il faudrait un effort pour mettre l'objet en mouvement et pour l'accélérer, à cause de l'inertie qu'il possède.

En mesurant avec précision la force exercée – par exemple en l'exerçant par l'intermédiaire d'un ressort et en mesurant l'allongement de celui-ci – on constaterait que la force f nécessaire pour obtenir une accélération donnée a est proportionnelle à la masse m. Si on double la masse, il faut une force double. Pour une masse donnée, la force doit être proportionnelle à l'accélération désirée. Mathématiquement, on peut exprimer cela par l'équation :

$$f = ma$$

Cette équation traduit la deuxième loi du mouvement, ou deuxième loi de Newton.

Or, comme l'a constaté Galilée, l'attraction terrestre accélère tous les corps, lourds ou légers, exactement de la même façon. (La résistance de l'air peut ralentir la chute d'objets très légers, mais dans le vide, comme on peut le montrer facilement, une plume tombera aussi vite qu'un morceau de plomb.) Si la deuxième loi est valable, il faut en conclure que l'attraction exercée par la Terre sur un objet lourd est plus grande que sur un objet léger, de manière à produire la même accélération. Pour accélérer une masse huit fois plus forte qu'une autre, par exemple, il faut une force huit fois plus grande. Il s'ensuit que l'attraction exercée par la Terre sur un objet doit être exactement proportionnelle à la masse de cet objet. (C'est pourquoi, en fait, on peut fort bien se servir du poids pour mesurer avec précision la masse.)

Newton a également énoncé une troisième loi du mouvement : à toute action correspond une réaction de même valeur et de direction opposée. Cette loi s'applique à la force. En d'autres termes, si la Terre tire sur la Lune avec une certaine force, la Lune tire sur la Terre exactement aussi fort. Si la masse de la Lune doublait, l'attraction que la Terre exerce sur elle doublerait aussi, en vertu de la deuxième loi. Par conséquent, en vertu de la troisième loi, la force exercée par la Lune sur la Terre doublerait aussi.

De la même façon, si la masse de la Terre doublait, la force qu'exerce sur elle la Lune devrait doubler, d'après la deuxième loi, et donc la force exercée par la Terre sur la Lune doublerait aussi, en vertu de la troisième loi.

Si les deux corps doublaient de masse, les forces auraient deux raisons de doubler et seraient multipliées par quatre.

A partir de raisonnements de ce type, Newton était forcé de conclure que la force gravitationnelle entre deux objets quelconques de l'Univers était proportionnelle au produit des masses de ces deux corps. Et, bien sûr, comme il l'avait déjà montré, cette force devait être inversement proportionnelle au carré de la distance entre les deux corps. C'est la loi de la gravitation universelle de Newton.

Si nous désignons la force gravitationnelle par *f*, les deux masses par *m* et *m'*, et la distance entre les deux corps par *d*, cette loi peut s'exprimer comme suit :

$$f = \frac{G\,m\,m'}{d^2}$$

G est la *constante de la gravitation*, dont la détermination a permis de « peser la Terre » (voir le chapitre 4). Newton avait fait l'hypothèse que cette constante devait avoir la même valeur à travers tout l'Univers. Au fil du temps, on s'est aperçu que les nouvelles planètes, inconnues à l'époque de Newton, avaient des mouvements conformes aux lois de Newton. Même la valse des étoiles doubles, prodigieusement éloignées, se mène à la cadence des lois qui expriment l'analyse newtonienne de l'Univers.

Tout ceci est sorti du nouveau point de vue quantitatif défendu par Galilée. Comme on peut le constater, la plus grande partie des mathématiques utilisées est très simple, et ce que nous avons cité ici s'enseigne au lycée.

En fait, ce qui était nécessaire pour introduire l'une des plus grandes révolutions scientifiques de tous les temps se réduit à :

- un ensemble d'observations simples que n'importe quel élève de lycée pourrait faire si on le guidait un peu ;
- un outillage mathématique du niveau du lycée ;
- le génie extraordinaire de Galilée et de Newton, qui leur a permis de faire les premiers ces observations et les généralisations mathématiques correspondantes.

La relativité

Les lois du mouvement, telles que Galilée et Newton les ont établies, reposent sur l'existence d'un déplacement absolu – c'est-à-dire un déplacement par rapport à un repère au repos. Mais tout ce que nous connaissons dans l'Univers est en mouvement : la Terre, le Soleil, notre galaxie, l'ensemble des galaxies... Où donc pouvons-nous trouver dans l'Univers quelque chose d'immobile, à quoi rapporter les mouvements que nous voulons étudier ?

L'EXPÉRIENCE DE MICHELSON-MORLEY

C'est ce type de réflexions qui a conduit à l'expérience de Michelson-Morley, qui à son tour a amené une révolution scientifique aussi importante, sous certains aspects, que celle de Galilée (voir le chapitre 8). Ici encore, les bases mathématiques sont assez simples.

L'expérience visait à détecter le mouvement de la Terre par rapport à l'*éther* qui était censé emplir tout l'espace et être immobile. Les idées traduites par cette expérience étaient les suivantes.

Supposons qu'on envoie un pinceau de lumière dans la direction du déplacement de la Terre par rapport à l'éther, et qu'à une certaine distance dans cette direction, un miroir fixé à la Terre renvoie vers la source le pinceau lumineux. Appelons c la vitesse de la lumière, v la vitesse de la Terre par rapport à l'éther, et d la distance entre la source et le miroir. La lumière part avec la vitesse $c + v$: sa vitesse propre, plus la vitesse de la Terre (pour ainsi dire, elle a « le vent dans le dos »). Elle met donc, pour atteindre le miroir, un temps égal à $d/c + v$.

Mais au retour, la situation n'est plus la même. La lumière a maintenant contre elle le « vent » de la vitesse de la Terre, et sa vitesse résultante est $c - v$. Le temps qu'elle met pour revenir à la source est $d/c - v$.

La durée totale de l'aller et retour est :

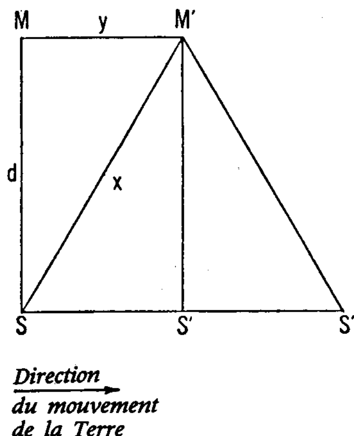
$$\frac{d}{c+v} + \frac{d}{c-v}$$

En combinant algébriquement les deux termes, on obtient :

$$\frac{d(c-v) + d(c+v)}{(c+v)\ (c-v)} = \frac{dc - dv + dc + dv}{c^2 - v^2} = \frac{2\,dc}{c^2 - v^2}$$

Supposons maintenant que le pinceau lumineux soit envoyé vers un miroir situé à la même distance, mais dans une direction perpendiculaire à celle du déplacement de la Terre par rapport à l'éther.

Le pinceau lumineux est pointé de S (la source) vers M (le miroir), qui sont séparés par la distance d. Mais pendant le temps mis par le pinceau



lumineux pour atteindre le miroir, le mouvement de la Terre a emmené le miroir M en M', de sorte que le chemin effectivement suivi par la lumière va de S à M'. Appelons cette distance x, et appelons y la distance MM' (voir schéma ci-dessus).

Pendant que la lumière parcourt la distance x à la vitesse c, le miroir parcourt la distance y à la vitesse (de la Terre) v. Puisque la lumière et

le miroir arrivent en même temps au point M′, les distances parcourues doivent être proportionnelles aux deux vitesses. Par conséquent :

$$\frac{y}{x} = \frac{v}{c} \quad \text{soit } y = \frac{vx}{c}$$

Maintenant, nous pouvons calculer x grâce au théorème de Pythagore, d'après lequel la somme des carrés des côtés de l'angle droit d'un triangle rectangle est égale au carré de l'hypothénuse. Dans le triangle SMM′, en remplaçant y par $\frac{vx}{c}$, on obtient :

$$x^2 = d^2 + \left(\frac{vx}{c}\right)^2$$

$$x^2 - \left(\frac{vx}{c}\right)^2 = d^2$$

$$x^2 - \frac{v^2x^2}{c^2} = d^2$$

$$\frac{c^2x^2 - v^2x^2}{c^2} = d^2$$

$$(c^2 - v^2)x^2 = d^2c^2$$

$$x^2 = \frac{d^2c^2}{c^2 - v^2}$$

$$x = \frac{dc}{\sqrt{c^2 - v^2}}$$

La lumière réfléchie sur le miroir en M′ revient à la source, qu'elle atteint en S. La distance S′S étant égale à SS′, M′S est égale à x. Le chemin total parcouru par la lumière est donc $2x$, soit : $2cd/\sqrt{c^2 - v^2}$.

Le temps mis par cette lumière pour parcourir cette distance (à la vitesse c) est :

$$\frac{2dc}{\sqrt{c^2 - v^2}} : c = \frac{2d}{\sqrt{c^2 - v^2}}$$

Comparons ce temps à celui que durerait l'aller-retour dans la direction du mouvement de la Terre. Pour cela, nous pouvons diviser le second par le premier :

$$\frac{2dc}{c^2 - v^2} : \frac{2d}{\sqrt{c^2 - v^2}} = \frac{2dc}{c^2 - v^2} \times \frac{\sqrt{c^2 - v^2}}{2d} = \frac{c}{\sqrt{c^2 - v^2}}$$

Cette expression se simplifie encore si on multiplie le numérateur et le dénominateur par $\sqrt{1/c^2}$ (qui est égal à $1/c$) :

$$\frac{c\sqrt{1/c^2}}{\sqrt{c^2 - v^2}\,\sqrt{1/c^2}} = \frac{1}{\sqrt{1 - v^2/c^2}}$$

Et voilà. C'est le rapport du temps mis par la lumière dans la direction du mouvement de la Terre au temps mis dans la direction perpendiculaire. Pour toute cette valeur de v supérieure à 0, ce rapport est plus grand que 1. Par conséquent, si la Terre se déplace bien par rapport à un éther immobile, la lumière devrait mettre plus longtemps pour faire un aller-retour dans le sens du mouvement de la Terre que pour faire un aller-retour dans la direction perpendiculaire. (En fait, le déplacement « parallèle » correspond à un temps maximal, et le temps « perpendiculaire » à un temps minimal.)

Michelson et Morley ont conçu leur expérience pour détecter la différence entre les temps mis par la lumière dans différentes directions. En dirigeant le pinceau de lumière dans toutes les directions, et en mesurant le temps mis pour un aller-retour grâce à leur interféromètre extrêmement précis, ils pensaient déceler des différences de vitesse apparente. La direction pour laquelle ils trouveraient une vitesse minimale de la lumière devrait coïncider avec celle du mouvement de la Terre, et la direction pour laquelle ils trouveraient une vitesse minimale devrait coïncider avec la perpendiculaire au mouvement de la Terre. A partir des valeurs de la vitesse de la lumière dans ces deux cas, on pourrait calculer la vitesse du mouvement de la Terre.

Seulement, l'expérience a donné exactement la même vitesse de la lumière dans toutes les directions ! Autrement dit, quel que soit le mouvement de la source, la vitesse de la lumière était toujours égale à c – en contradiction évidente avec les lois du mouvement énoncées par Newton. En tentant de mesurer le mouvement absolu de la Terre, Michelson et Morley avaient ainsi non seulement introduit des doutes sérieux quant à l'existence de l'éther, mais aussi mis en cause l'idée de mouvement absolu et de repos absolu, et les fondements mêmes de l'univers newtonien.

L'ÉQUATION DE FITZGERALD

Le physicien irlandais G.F. FitzGerald a découvert une manière de « sauver les meubles ». Il a émis l'idée que les objets se raccourcissaient dans la direction de leur mouvement, dans un rapport égal à $\sqrt{1 - v^2/c^2}$. Ainsi :

$$L' = L \sqrt{1 - v^2/c^2}$$

L' étant la longueur d'un corps en mouvement, suivant la direction de ce mouvement, et L la valeur qu'aurait cette longueur au repos.

Ce raccourcissement compenserait exactement le rapport des deux vitesses de la lumière dans l'expérience de Michelson et Morley. Ce rapport deviendrait égal à 1 : pour nos instruments (et nos organes) raccourcis convenablement, la vitesse de la lumière semblerait la même dans toutes les directions, quel que soit le mouvement de la source par rapport à l'éther.

Dans des conditions normales, ce raccourcissement est très faible. Même si un corps se déplaçait à une vitesse égale au dixième de celle de la lumière (à 30 000 km/s), sa longueur ne diminuerait que très peu, d'après l'équation de FitzGerald :

$$L' = L \sqrt{1 - \left(\frac{0{,}1}{1}\right)^2} = L \sqrt{1 - 0{,}01}$$

$$L' = L \sqrt{0{,}99}$$

Ainsi, L' vaut à peu près 0,995 L : le raccourcissement n'est que de 0,5 %. Des corps se déplaçant à des vitesses de cet ordre ne se rencontrent que dans le domaine des particules subatomiques. Quant à un avion volant à 3 000 km/h, son raccourcissement est minuscule, comme on peut le calculer facilement.

A quelle vitesse un objet sera-t-il raccourci de moitié ? Remplaçons L' par 1/2 L dans l'équation de FitzGerald :

$$\frac{L}{2} = L\sqrt{1 - v^2/c^2}$$

Soit, en divisant par L :

$$\frac{1}{2} = \sqrt{1 - v^2/c^2}$$

Elevons au carré les deux membres :

$$\frac{1}{4} = 1 - v^2/c^2$$

$$v^2/c^2 = \frac{3}{4}$$

$$v = \sqrt{\frac{3}{4}}\, c = 0{,}866\, c$$

La vitesse de la lumière dans le vide valant 300 000 km/s, la vitesse à laquelle un objet est raccourci de moitié est 0,866 fois 300 000, soit environ 260 000 km/s.

Si un objet se déplace à la vitesse de la lumière, de sorte que sa vitesse est $v = c$, l'équation de FitzGerald devient :

$$L' = L\sqrt{1 - c^2/c^2} = L\sqrt{0} = 0$$

Ainsi, à la vitesse de la lumière, la longueur suivant la direction du mouvement s'annule. Il semble donc qu'une vitesse supérieure à celle de la lumière soit impossible.

L'ÉQUATION DE LORENTZ

Quelques années après que FitzGerald eut proposé son équation, on a découvert l'électron, et les physiciens ont commencé à s'intéresser aux particules chargées. Lorentz a élaboré une théorie suivant laquelle la masse d'une particule de charge donnée serait inversement proportionnelle à son rayon. En d'autres termes, plus la charge est concentrée dans un petit volume, plus la masse est grande.

Or, si une particule est raccourcie par son mouvement, son rayon suivant la direction du mouvement diminue suivant l'équation de FitzGerald. En remplaçant dans cette équation L et L' par R et R', on obtient :

$$R' = R\sqrt{1 - \frac{v^2}{c^2}} \text{ soit } R'/R = \sqrt{1 - v^2/c^2}$$

Si la masse de la particule est inversement proportionnelle à son rayon :

$$\frac{R'}{R} = \frac{M}{M'}$$

M étant la masse de la particule au repos et M' sa masse en mouvement. En remplaçant dans l'équation précédente R'/R par M/M' ; on obtient :

$$M/M' = \sqrt{1 - v^2/c^2} \text{ soit } M' = \frac{M}{\sqrt{1 - v^2/c^2}}$$

L'équation de Lorentz se manipule de la même façon que celle de FitzGerald. Par exemple, pour une particule se déplaçant à 30 000 km/s (un dixième de la vitesse de la lumière), la masse M' serait supérieure de 0,5 % à la masse M au repos. A 260 000 km/s, la masse apparente serait le double de la masse au repos.

Finalement, pour une particule se déplaçant à la vitesse de la lumière, $v = c$, et l'équation de Lorentz devient :

$$M' = \frac{M}{\sqrt{1 - c^2/c^2}} = \frac{M}{O}$$

Or, quand le dénominateur d'une fraction *tend vers zéro*, la fraction devient infiniment grande. Autrement dit, il semblerait que la masse d'un objet voyageant à une vitesse approchant celle de la lumière devienne infiniment grande. Une fois encore, la vitesse de la lumière semble bien être le maximum possible.

Tout ceci a conduit Einstein à refondre les lois du mouvement et de la gravitation. Dans son univers, les résultats de l'expérience de Michelson-Morley sont tout à fait normaux.

Mais nous ne sommes pas au bout de nos découvertes.

Remarquons que l'équation de Lorentz attribue à la masse M au repos une valeur non nulle. C'est vrai de la plupart des particules qui nous sont familières, et de tous les objets, des atomes aux étoiles, qui sont constitués de ces particules. Mais pour les neutrinos et les antineutrinos, la masse au repos M est nulle. C'est également le cas des photons.

De telles particules, si elles sont dans le vide, s'y déplacent avec la même vitesse que celle de la lumière. Sitôt créées, elles filent à cette vitesse, sans période d'accélération appréciable.

On peut se demander comment il est possible de parler de « masse au repos » pour le photon ou le neutrino, s'ils ne sont jamais au repos, mais n'existent qu'en se déplaçant à 300 000 km/s (dans le vide). Pour cette raison, les physiciens Olexa-Myron Bilaniuk et Ennackal Chandy George Sudarshan ont proposé d'appeler M *masse propre* plutôt que « masse au repos ». Pour une particule de masse positive, la masse propre est celle que l'on mesure quand la particule est immobile par rapport aux instruments de mesure et à l'observateur. Pour une particule de masse nulle, on obtient la masse propre par un raisonnement indirect. Bilaniuk et Sudarshan ont suggéré d'appeler *luxons* (d'un mot latin signifiant « lumière ») les objets de masse propre nulle, qui se déplacent à la vitesse de la lumière, et *tardyons* les particules de masse propre positive, qui se déplacent à des vitesses inférieures à celle de la lumière.

En 1962, Bilaniuk et Sudarshan ont commencé à se poser des questions sur les vitesses supérieures à celle de la lumière. Une particule se déplaçant à une telle vitesse aurait une masse imaginaire. C'est-à-dire que cette masse serait égale à un nombre ordinaire multiplié par la racine carrée de (–1).

Imaginons, par exemple, une particule se déplaçant deux fois plus vite que la lumière ($v = 2c$ dans l'équation de Lorentz) :

$$M' = \frac{M}{\sqrt{1-(2c)^2/c^2}} = \frac{M}{\sqrt{1-4c^2/c^2}} = \frac{M}{\sqrt{-3}}$$

Ainsi, la masse en mouvement serait le produit de la masse propre par un nombre imaginaire, d'où on peut déduire que les particules se déplaçant à une vitesse supérieure à celle de la lumière doivent avoir une masse propre imaginaire. Cela signifie-t-il que de telles particules ne peuvent pas exister ?

Pas nécessairement. En admettant l'existence de masses propres imaginaires, nous pouvons appliquer à de telles particules toutes les équations de la théorie de la relativité restreinte d'Einstein. Mais de telles particules ont une propriété apparemment paradoxale : plus elles vont lentement, plus elles possèdent d'énergie. Ceci est l'inverse de ce qui se passe dans notre univers, et c'est peut-être là la signification des masses propres imaginaires. Une particule de masse imaginaire accélère en présence d'une résistance, et ralentit quand on la pousse en avant. A mesure que son énergie diminue, sa vitesse augmente, jusqu'à ce qu'elle arrive, avec une énergie nulle, à la vitesse infinie. A mesure que son énergie augmente, elle se déplace de plus en plus lentement, jusqu'à approcher la vitesse de la lumière, avec une énergie infinie.

Le physicien américain Gerald Feinberg a donné à ces particules le nom de *tachyons* (d'un mot grec signifiant « rapide »).

Nous pouvons donc imaginer deux sortes d'univers. La première, correspondant à notre univers, est constituée de tardyons, particules qui vont moins vite que la lumière et peuvent accélérer jusqu'à s'en approcher, tandis que leur énergie augmente. L'autre catégorie d'univers est composée de tachyons, particules plus rapides que la lumière, qui peuvent ralentir jusqu'à sa vitesse, tandis que leur énergie augmente. Entre les deux se trouve le *mur des luxons*, où se trouvent les particules qui vont exactement aussi vite que la lumière. Ce mur est mitoyen entre les deux univers.

Si un tachyon possède assez d'énergie, et par conséquent se déplace assez lentement, il peut rester à la même place assez longtemps pour engendrer une explosion de photons détectables. (Les tachyons laisseraient une traînée de photons même dans le vide, à cause d'une sorte d'effet Tchérenkov.) Les physiciens guettent une telle explosion, mais il y a peu de chances d'avoir un instrument à l'endroit exact où l'une de ces explosions (probablement très rares) pourrait se produire, pendant un milliardième de seconde ou moins.

Pour certains physiciens, « tout ce qui n'est pas interdit est obligatoire ». Autrement dit, tout phénomène qui ne viole aucune loi de conservation *doit* tôt ou tard se produire ; si les tachyons ne violent pas la relativité restreinte, ils *doivent* exister. Toutefois, les adeptes de cette théorie accueilleraient avec plaisir (et peut-être avec soulagement) la moindre preuve expérimentale que ces tachyons non interdits existent bien. Jusqu'ici, ils n'en ont pas eu le plaisir.

L'ÉQUATION D'EINSTEIN

Einstein a tiré de l'équation de Lorentz une conséquence qui est devenue l'équation la plus célèbre peut-être de tous les temps.

On peut écrire l'équation de Lorentz sous la forme :

$$M' = M\,(1 - v^2/c^2)^{-\frac{1}{2}}$$

en notant $1/\sqrt{x}$ par $x^{-\frac{1}{2}}$. Ceci met l'équation sous une forme que l'on peut développer en une série de termes, comme l'a montré, par un curieux hasard, Newton lui-même. C'est ce que l'on appelle la formule du binôme.

Le nombre des termes du développement est infini, mais comme chacun d'eux est plus petit que le précédent, en prenant seulement les deux premiers, on obtient un résultat approximativement correct, la somme de tous les autres étant négligeable. Le développement devient ainsi :

$$\left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}} = 1 + \frac{\frac{1}{2}\,v^2}{c^2} + \dots$$

En substituant dans l'équation de Lorentz, on obtient :

$$M' = M\left(1 + \frac{\frac{1}{2}\,v^2}{c^2}\right) = M + \frac{\frac{1}{2}\,Mv^2}{c^2}$$

Or, en physique classique, l'expression $\frac{1}{2}\,Mv^2$ représente l'énergie d'un corps en mouvement. Si nous notons l'énergie par la lettre e, l'équation ci-dessus devient :

$$M' = M + e/c^2 \text{ soit } M' - M = e/c^2$$

L'augmentation de masse due au mouvement ($M' - M$) peut être notée m. Ainsi :

$$m = e/c^2 \text{ soit } e = mc^2$$

C'est cette équation qui a indiqué pour la première fois que la masse est une forme de l'énergie. Einstein a été plus loin, en montrant que l'équation s'applique à toute masse, et pas seulement à l'augmentation de masse due au mouvement.

Ici encore, la plus grande partie des mathématiques nécessaires est du niveau du lycée. Et pourtant, le monde y a trouvé une vision de l'Univers encore plus large que celle de Newton, et aussi un certain nombre d'applications concrètes, comme le réacteur nucléaire et la bombe atomique.

Bibliographie

Généralités

ALONSO, M. et FINN, E. J. *Physique générale* (2e éd., 2 vol.). Paris : InterÉditions, 1986.

Annuaires du Bureau des Longitudes. *Encyclopédie scientifique de l'univers* (4 vol.). Paris : Gauthier-Villars, 1980, 1981.

ASIMOV, I. *Asimov'Biographical Encyclopedia of Science and Technology.* New York : Doubleday, 1982.

ASIMOV, I. *La conquête du savoir.* Verviers : Marabout, 1984.

ASIMOV, I. *Le monstre subatomique.* Paris : Londreys, 1985.

ASIMOV, I. *X comme inconnu.* Paris : Londreys, 1984.

BERKELEY. *Cours de physique* (5e éd., 5 vol.). Paris : Armand Colin, 1984.

BLANC, M. éd. *L'état des sciences.* Paris : La Découverte, 1983.

CRONIN, V. *La Terre, le cosmos et l'homme.* Paris : Denoël, 1981.

FEYNMAN, R. *La nature de la physique.* Paris : Le Seuil/Points Science, 1980.

FEYNMAN, R., LEIGHTON, R. et SANDS, M. *Le cours de physique de Feynman* (5 vol.). Paris : InterÉditions, 1979.

LANDAU, L. et KITAIGORODSKI, A. *La physique à la portée de tous* (4 vol.). Moscou : Mir, 1984.

RUFFIE, J. *De la biologie à la culture.* Paris : Flammarion, 1976.

RUSSO, F. *Pour une bibliothèque scientifique.* Paris : Le Seuil/Points Science, 1972.

SAGAN, C. *Cosmos.* Paris : Mazarine, 1981.

Chapitre 1 Qu'est-ce que la science ?

BERNAL, J. D. *Science in History.* New York : Hawthorn Books, 1965.

CLAGETT, M. *Greek Science in Antiquity.* New York : Abelard-Schuman, 1955.

Collectif. *La Recherche en histoire des sciences.* Paris : Le Seuil/Points Science, 1983.

COUDERC, P. *Histoire de l'astronomie classique.* Paris : P.U.F./Que sais-je ?, 1982.

CROMBIE. *Histoire des sciences de Saint Augustin à Galilée (400-1650)* (2 vol.). Paris : P.U.F., 1959.

DAMPIER, SIR W. C. *Histoire de la science.* Paris : Payot, 1951.

DREYER, J. L. E. *A History of Astronomy from Thales to Kepler.* New York : Dover Publications, 1953.

EINSTEIN, A. et INFELD, L. *L'évolution des idées en physique.* Paris : Payot, 1981.

FARRINGTON, B. *La science dans l'antiquité.* Paris : Payot, 1967.

RONIN, C. A. *Science : Its History and Development among the World's Cultures.* New York : Facts on File Publications, 1982.

ROSMORDUC, J. *Une histoire de la physique et de la chimie : de Thalès à Einstein.* Paris : Le Seuil/Points Science, 1985.

ROUSSEAU, P. *Histoire de la science.* Paris : Fayard, 1965.

SEGRE, E. *Les physiciens modernes et leurs découvertes.* Paris : Fayard, 1984.

TATON, R. *Histoire générale des sciences* (4 vol.). Paris : P.U.F., 1966 à 1983.

Chapitre 2 L'Univers

ABELL, G. O. *Exploration of the Universe.* Philadelphie : Saunders College Publishing, 1982.

ASIMOV, I. *Myriades.* Paris : Londreys, 1984.

ASIMOV, I. *Supernova.* Paris : Payot, 1986.

AUDOUZE, J. *Aujourd'hui l'Univers.* Paris : Belfond, 1984.

BOISCHOT, A. *La radioastronomie.* Paris : P.U.F./Que sais-je ?, 1971.

Collectif. *La Recherche en astrophysique.* Paris : Le Seuil/Points Science, 1977.

Collectif. *Vie et mort des étoiles.* Paris : Belin/Bibliothèque pour la Science, 1985.

FERRIST, T. *Galaxies.* Paris : Mazarine, 1983.

FLAMMARION, C. *Astronomie populaire.* Paris : Flammarion, 1880 (rééd. 1980).

GATTY, B. *Hier, l'Univers.* Paris : Hermann, 1985.

ISLAM, J. *Le destin ultime de l'Univers.* Paris : Belfond, 1984.

JASTROW, R. *Des astres, de la vie et des hommes.* Paris : Le Seuil/Points Science, 1975.

KIPPENHAM, R. *100 Billions Suns.* New York : Basic Books, 1983.

MOORE, P. *Éléments d'astronomie.* Paris : Dunod/Science Poche, 1970.

PECKER, J.-C. éd. *Astronomie Flammarion.* Paris : Flammarion, 1985.

REEVES, H. *Patience dans l'azur.* Paris : Le Seuil, 1981.

REEVES, H. *Poussières d'étoiles.* Paris : Le Seuil/Science ouverte, 1984.

SAGAN, C. *Cosmic Connection ou l'appel des étoiles.* Paris : Le Seuil/Points Science, 1978.

SILK, J. et BARROW, J. D. *La main gauche de la création.* Paris : Londreys, 1985.

VERON, P. *Les quasars.* Paris : P.U.F./Que sais-je ?, 1974.

WEINBERG, S. *Les trois premières minutes de l'Univers.* Paris : Le Seuil, 1978.

Chapitre 3 Le système solaire

BEATTY, J. K., O'LEARY, B. et CHAIKIN, A. eds. *The New Solar System.* Cambridge, Mass : Sky Publishing, et Cambridge, Angleterre : Cambridge University Press, 1981.

Collectif. *Le système solaire.* Paris : Belin/Bibliothèque pour la Science, 1982.
MOORE, P. et MASON, J. *Le retour de la comète de Halley.* Paris : Londreys, 1985.
PECKER, J. C. *Sous l'étoile soleil.* Paris : Fayard, 1984.
RYAN, P. et PESEK, L. *Solar System.* New York : Viking Press, 1978.
SAGAN, C. *Comètes.* Paris : Calmann-Lévy, 1985.

Chapitre 4 La Terre

ALBAREDE, F. et CONDOMINES, M. *La géochimie.* Paris : P.U.F./Que sais-je ?, 1976.
ALLEGRE, C. *L'écume de la Terre.* Paris : Fayard, 1983.
CAILLEUX, A. *La géologie.* Paris : P.U.F./Que sais-je ?, 1982.
CAILLEUX, A. *Histoire de la géologie.* Paris : P.U.F./Que sais-je ?, 1968.
CARRE, F. *Les océans.* Paris : P.U.F./Que sais-je ?, 1983.
Collectif. *La dérive des continents.* Paris : Belin/Bibliothèque pour la Science, 1982.
Collectif. *Les tremblements de terre.* Belin/Bibliothèque pour la Science, 1982.
Collectif. *Les volcans.* Paris : Belin/Bibliothèque pour la Science, 1984.
Collectif. *La recherche sur les énergies nouvelles.* Paris : Le Seuil/Points Science, 1980.
HALLAM, A. *Une révolution dans les sciences de la Terre : de la dérive des continents à la tectonique des plaques.* Paris : Le Seuil/Points Science, 1976.
JACKSON, D. D. *Underground Worlds.* Alexandria, Va. : Time-Life Books, 1982.
MORET, L. *Précis de géologie.* Paris : Masson, 1967.
NOVIKOV, E. *La planète des énigmes.* Moscou : Mir, 1978.
ROMANOVSKY, V. *Physique de l'océan.* Paris : Le Seuil, 1966.
ROUBAULT, M. et COPPENS, R. *La dérive des continents.* (2e éd.). Paris : P.U.F./Que sais-je ?, 1979.
TAZIEFF, H. *Les volcans et la dérive des continents.* Paris : P.U.F., 1984.

Chapitre 5 L'atmosphère

ALLEN, J. *Aller-retour pour l'espace.* Paris : Mazarine, 1984.
Collectif. *Les phénomènes naturels.* Paris : Belin/Bibliothèque pour la Science, 1978.
DELOBEAU, F. *L'environnement de la Terre.* Paris : P.U.F., 1968.
FLAMMARION, C. *L'atmosphère.* Paris : Hachette, 1872.
LAMING, L. *L'astronautique.* Paris : P.U.F./Que sais-je ?, 1971.
LEVASSEUR-REGOURD, A. C. *L'atmosphère et ses phénomènes.* De Vecchi, 1980.
MARTIN, C.-N. *Les satellites artificiels.* Paris : P.U.F./Que sais-je ?, 1972.
MEZENTSEV, M. *Phénomènes étranges dans l'atmosphère et sur la Terre.* Moscou : Mir, 1970.

PICCARD, A. *Au seuil du cosmos.* Paris : Denoël, 1964.

VIAUL, A. *La météorologie.* Paris : P.U.F./Que sais-je ?, 1978.

VIRGATCHIK, I. *Le guide Marabout de la météorologie, des phénomènes atmosphériques et des microclimats.* Verviers : Marabout, 1981.

YOUNG, L. B. *Earth's Aura.* New York : Alfred A. Knopf, 1977.

Chapitre 6 Les éléments

ASIMOV, I. *A Short History of Chemistry.* New York : Doubleday, 1965.

BORDRY, M. et RADVANYI, P. *La radioactivité artificielle et son histoire.* Paris : Le Seuil/Points Science, 1984.

CHAMPETIER, G. *La chimie générale.* Paris : P.U.F./Que sais-je ?, 1979.

DAGOGNET, F. *Tableaux et langages de la chimie.* Paris : Le Seuil, 1969.

DUCROCQ, A. *Les éléments au pouvoir.* Paris : Julliard, 1976.

EUDIER, M. *Les alliages métalliques.* Paris : P.U.F./Que sais-je ?, 1976.

HOCART, R. *Les cristaux.* Paris : P.U.F./Que sais-je ?, 1974.

HOCHEID, R. *La métallurgie.* Paris : P.U.F./Que sais-je ?, 1980.

PAULING, L. *Chimie générale.* Paris : Dunod, 1966.

REID, R. *Marie Curie.* Paris : Le Seuil/Points Science, 1983.

SLADE, E. *Une métallurgie nouvelle.* Paris : Larousse/Technique d'aujourd'hui, 1971.

WOJTKOWIAK, B. *Histoire de la chimie.* Paris : Lavoisier, 1984.

Chapitre 7 Les particules

ALFREN, H. *Worlds Antiworlds.* San Francisco : W. H. Freeman, 1966.

ASIMOV, I. *Une particule fantôme, le neutrino.* Paris : Dunod, 1970.

BLANC, D. *La chimie nucléaire.* Paris : P.U.F./Que sais-je ?, 1986.

BLANC, D. *La physique nucléaire.* Paris : P.U.F./Que sais-je ?, 1984.

Collectif. *Les particules élémentaires.* Paris : Belin/Bibliothèque pour la Science, 1983.

Collectif. *La Recherche en physique nucléaire.* Paris : Le Seuil/Points Science, 1983.

FEINBERG, G. *What is the World Made Of?* Garden City, New York : Anchor Press/Doubleday, 1977.

GARDNER, M. *L'univers ambidextre.* Paris : Le Seuil, 1986.

KAHAN, T. *Les particules élémentaires.* Paris : P.U.F./Que sais-je ?, 1977.

LEFORT, M. *Les radiations nucléaires.* Paris : P.U.F./Que sais-je ?, 1980.

NOËL, E. éd. *La matière aujourd'hui.* Paris : Le Seuil/Points Science, 1981.

WEINBERG, S. *Le monde des particules.* Paris : Belin, 1983.

Chapitre 8 Les ondes

ATKINS, P. W. *Chaleur et désordre.* Paris : Belin, 1986.
BERKELEY. *Cours de physique* (5e éd., 5 vol.). Paris : Armand Colin, 1984.
BORY, C. *La thermodynamique.* Paris : P.U.F./Que sais-je ?, 1981.
DELIGEORGES, S. éd. *Le monde quantique.* Paris : Le Seuil/Points Science, 1985.
EINSTEIN, A. *La relativité.* Paris : Payot, 1983.
GARDNER, M. *La relativité pour tous.* Paris : Dunod, 1969.
HOFFMANN, B. et PATY, M. *L'étrange histoire des quanta.* Paris : Le Seuil/Points Science, 1981.
LEVY-LEBLOND, J.-M. et BALIBAR, F. *Quantique.* Paris : InterÉditions, 1984.
MAITTE, B. *La lumière.* Paris : Le Seuil/Points Science, 1981.
NOËL, E. éd. *L'espace et le temps aujourd'hui.* Paris : Le Seuil/Points Science, 1983.
ORTOLI, S. et PHARABOD, J.-P. *Le cantique des quantiques : le monde existe-t-il ?* Paris : La Découverte, 1984.
PAGELS, H. *L'univers quantique.* Paris : InterÉditions, 1985.
SMITH, J. *Introduction à la relativité.* Paris : InterÉditions, 1981.

Chapitre 9 La machine

CLARKE, D. éd. *The Encyclopedia of How it Works.* New York : A & W Publishers, 1977.
Collectif. *Histoire de machines.* Paris : Belin/Bibliothèque pour la Science, 1982.
Collectif. *Inventions et découvertes.* Paris : Instant, 1985.
DAUMAS, M. *Les grandes étapes du progrès technique.* Paris : P.U.F./Que sais-je ?, 1981.
DAUMAS, M. éd. *Histoire générale des techniques* (5 vol.). Paris : P.U.F., 1962 à 1979.
DAVID, P. *L'électronique.* Paris : P.U.F./Que sais-je ?, 1983.
DEZOTEUX, J. *Les transistors.* Paris : P.U.F./Que sais-je ?, 1980.
DYAN, B. et CHARLES, G. éds. *Guide des technologies de l'information.* Paris : Autrement, 1984.
GILLE, B. *Les ingénieurs de la Renaissance.* Paris : Le Seuil/Points Science, 1978.
GILLE, B. *Les mécaniciens grecs.* Paris : Le Seuil/Science ouverte, 1980.
GUILLEN, R. *Les semiconducteurs et leurs applications.* Paris : P.U.F./Que sais-je ?, 1978.
HERLEA, A. *Les moteurs.* Paris : P.U.F./Que sais-je ?, 1985.
KLEMM, F. *Histoire des techniques.* Paris : Payot, 1966.
LION, P. et CHAUSSIN, C. tr. *Comment ça marche ?* Paris : Dunod, 1971.
MCCAIG, M. *Les aimants.* Paris : Dunod/Science Poche, 1970.
NATTA, G. et PASQUON, I. *Les principes de la chimie industrielle.* Paris : Gauthier-Villars, 1969.
PIERCE, J.-R. *Symboles, signaux et bruits.* Paris : Masson, 1966.
SANDORI, P. *Petite logique des forces : constructions et machines.* Paris : Le Seuil/Points Science, 1983.

SOBOLEV, N. *Les lasers, réalisations et perspectives.* Moscou : Mir, 1973.
TUMA, J. *Encyclopédie illustrée des transports.* Paris : Gründ, 1982.

Chapitre 10 Le réacteur

AUDIBERT, P. et ROUARD, D. *Les énergies du Soleil.* Paris : Le Seuil/Points Science, 1978.
BIZEC, R. éd. *La recherche sur les énergies nouvelles.* Paris : Le Seuil/Points Science, 1980.
BLANC, D. *Les réacteurs atomiques.* Paris : P.U.F./Que sais-je ?, 1986.
C.F.D.T., E. A. *Le dossier électronucléaire.* Paris : P.U.F./Que sais-je ?, 1980.
C.F.D.T., Groupe Confédéral Énergie. *Le dossier de l'énergie.* Paris : Le Seuil/Points Science, 1984.
CHANT, C ; HOGG, I. et VICTOR, J.-C. *La bombe : armes et scénarios nucléaires.* Paris : Autrement, 1983.
DUPUY, G. *Radioactivité : énergie nucléaire.* Paris : P.U.F./Que sais-je ?, 1982.
ETIEVANT, C. *L'énergie thermonucléaire.* Paris : P.U.F./Que sais-je ?, 1976.
GALLE, P. *Toxiques nucléaires.* Paris : Masson, 1982.
GUERON, J. *L'énergie nucléaire.* Paris : P.U.F./Que sais-je ?, 1982.
GUGLIEMO, R. et GUGLIEMO, G. *L'énergie nucléaire et les autres sources d'énergie.* Paris : Hatier, 1978.
WEART, S. *La grande aventure des atomistes français.* Paris : Fayard, 1980.

Chapitre 11 La molécule

GAUTIER, J.-A. et MIOCQUE, M. *Précis de chimie organique* (2 vol.). Paris : Masson, 1968 et 1969.
GEISSMAN, T. A. *Principes de chimie organique.* Paris : Dunod, 1965.
JACQUES, J. *Les confessions d'un chimiste ordinaire.* Paris : Le Seuil/Science ouverte, 1981.
PAULING, L. *The Nature of the Chemical Bond* (3e éd.). Ithaca, N.Y. : Cornell University Press, 1960.
PAULING, L. et HAYWARD, R. *The Architecture of Molecules.* San Francisco : W. H. Freeman, 1964.
PULLMANN, B. *La structure moléculaire.* Paris : P.U.F./Que sais-je ?, 1973.
ROBERTS, J. D. et CASERIO, M. C. *Chimie organique moderne.* Paris : InterÉditions, 1968.
SALEM, L. *Molécule la merveilleuse.* Paris : InterÉditions, 1979.

Chapitre 12 Les protéines

APELGOT, S. *Introduction à l'emploi des radioéléments en biologie et en biochimie.* Paris : Masson, 1981.

BALDWIN, E. *Dynamic Aspects of Biochemistry* (5e éd.). New York : Cambridge University Press, 1967.

BALDWIN, E. *The Nature of Biochemistry.* New York : Cambridge University Press, 1962.

HARPER, H. *Précis de biochimie.* Presses Universitaires de Laval, 1982.

LAVOLLAY, J. *La chimie des êtres vivants.* Paris : P.U.F./Que sais-je ?, 1980.

LEHNINGER, A. L. *Biochimie.* Paris : Flammarion Médecine-Sciences, 1977.

LEHNINGER, A. L. *Bioénergétique.* Paris : InterÉditions, 1972.

RELYVELD, E. H. et CHERMAN, J. C. *Les protéines.* Paris : P.U.F./Que sais-je ?, 1980.

STOLKOWSKI, J. *Les enzymes.* Paris : P.U.F./Que sais-je ?, 1983.

Chapitre 13 La cellule

CARR, D. E. *The Deadly Feast of Life.* Garden City, New York : Doubleday, 1971.

Collectif. *Des gènes aux protéines.* Paris : Belin/Bibliothèque pour la Science, 1985.

Collectif. *Hérédité et manipulations génétiques.* Paris : Belin/Bibliothèque pour la Science, 1982.

Collectif. *La Recherche en biologie moléculaire.* Paris : Le Seuil/Points Science, 1975.

FEINBERG, G. et SHAPIRO, R. *Life Beyond Earth.* New York : William Morrow, 1980.

FIRKET, H. *La cellule vivante.* Paris : P.U.F./Que sais-je ?, 1982.

HOAGLAND, M. *La cellule insolite.* Paris : InterÉditions, 1981.

JACOB, F. *La logique du vivant.* Paris : Gallimard, 1970.

LWOFF, A. *L'ordre biologique.* Paris : Laffont, 1969.

MONOD, J. *Le hasard et la nécessité : essai sur la philosophie naturelle de la biologie moderne.* Paris : Le Seuil/Points Science, 1973.

OPARINE, A. *L'origine de la vie sur la Terre.* Paris : Masson, 1965.

ROSNAY, J. *Les origines de la vie.* Paris : Le Seuil/Points Science, 1975.

WATSON, J. *La double hélice.* Paris : Laffont, 1968.

WATSON, J. *Biologie moléculaire du gène.* Paris : InterÉditions, 1968.

Chapitre 14 Les micro-organismes

AYLESWORTH, T. G. *Le monde des microbes.* Paris : Flammarion, 1975.

BURDIN, J.-C. et LAVERGNE, E. *Les bactéries.* Paris : P.U.F./Que sais-je ?, 1978.

Collectif. *La Recherche sur le cancer.* Paris : Le Seuil/Points Science, 1982.

CURTIS, H. *The Viruses.* Garden City, New York : Natural History Press, 1965.

DELAUNAY, A. *Présence de Pasteur,* Paris : Fayard, 1973.

FOUGEREAU, M. *L'immunologie.* Paris : P.U.F./Que sais-je ?, 1979.
LATOUR, B. *Les microbes : guerre et paix.* Paris : Métaillé, 1984.
LEPINE, P. *Les virus.* Paris : P.U.F./Que sais-je ?, 1984.
McGRADY, P. *The Savage Cell.* New York : Basic Books, 1964.
MIZELS, S. et JARET, P. *Notre corps se défend.* Paris : Payot, 1986.
NILSON, L. *Le corps victorieux.* Paris : Le Chêne, 1986.

Chapitre 15 Le corps

BARRAU, J. *Les hommes et leurs aliments, esquisse d'une histoire écologique et ethnologique de l'alimentation humaine.* Paris : Messidor/Temps Actuels, 1983.
JACOB, A. *La nutrition.* Paris : P.U.F./Que sais-je ?, 1981.
JACOTOT, B. et al. *Nutrition et alimentation.* Paris : Masson, 1983.
KRAUSE, M. V. et HUNSHER, M. A. *Nutrition et diéthothérapie.* Montréal : HRW Ltée, 1978.
NOËL, J. *Le corps humain.* Paris : Gamma, 1976.
REY, P. *Les hormones.* Paris : P.U.F./Que sais-je ?, 1975.
SMITH, A. *The Body.* London : George Allen & Unwin, 1968.
SOLOMON, E. P. et DAVIS, P. W. *Anatomie et physiologie humaine.* Paris : McGraw-Hill, 1981.
TANNAHILL, R. *Food in History.* New York : Stein & Day, 1973.
TREMOLIERE, J. ; SERVILLE, Y. et JACQUOT, R. *Manuel élémentaire d'alimentation humaine.* Paris : Édition Sociale Française, 1967.
WILLIAMS, S. R. *Essentials of Nutrition and Diet Therapy.* St. Louis : C. V. Mosby, 1974.

Chapitre 16 Les espèces

BOULE, M. et MARCELLIN. *Les hommes fossiles : éléments de paléontologie humaine.* Paris : Masson, 1946.
CALVIN, M. *Chemical Evolution.* New York : Oxford University Press, 1969.
CAMPBELL, B. *Human Evolution* (2e éd.). Chicago : Aldine Publishing, 1974.
Collectif. *L'aube de l'humanité.* Paris : Belin/Bibliothèque pour la Science, 1983.
Collectif. *L'évolution.* Paris : Belin/Bibliothèque pour la Science, 1979.
Collectif. *La Recherche sur la génétique et l'hérédité.* Paris : Le Seuil/Points Science, 1985.
COPPENS, Y. *Le singe, l'Afrique et l'homme.* Paris : Fayard, 1983.
DE BELL, G. *The Environmental Handbook.* New York : Ballantine Books, 1970.
EHRLICH, P. et EHRLICH, A. *Extinction.* New York : Random House, 1981.
GATTY, B. *L'œuf du vivant.* Paris : Hermann, 1985.
GOULD, S. J. *Darwin et les grandes énigmes de la vie.* Paris : Le Seuil/Points Science, 1984.
GOULD, S. J. *Le pouce du panda.* Paris : Le Livre de Poche, 1986.

JACQUARD, A. *Éloge de la différence : la génétique et les hommes.* Paris : Le Seuil/Points Science, 1981.

NOËL, E. éd. *Le darwinisme aujourd'hui.* Paris : Le Seuil/Points Science, 1979.

ROSTAND, J. *L'homme.* Paris : Gallimard/Idées, 1970.

ROSTAND, J. *Peut-on modifier l'homme ?* Paris : Gallimard, 1956.

SIMPSON, G. G. *L'évolution et sa signification.* Paris : Payot, 1951.

TAYLOR, G. R. *The Doomsday Book.* New York : World Publishing, 1970.

Chapitre 17 L'esprit

BONNET, A. *L'intelligence artificielle, promesses et réalités.* Paris : InterÉditions, 1984.

CHANGEUX, J.-P. *L'homme neuronal.* Paris : Fayard, 1983.

Collectif. *Le cerveau.* Paris : Belin/Bibliothèque pour la Science, 1982.

Collectif. *La Recherche en éthologie : les comportements animaux et humains.* Paris : Le Seuil/Points Science, 1979.

Collectif. *La Recherche en neurobiologie.* Paris : Le Seuil/Points Science, 1977.

CYRULNIK, B. *Mémoire de singe et parole d'hommes.* Paris : Hachette-Pluriel, 1983.

DEMARNE, P. et ROUQUEROL, M. *Les ordinateurs électroniques.* Paris : P.U.F./Que sais-je ?, 1985.

EVANS, C. *Les géants minuscules.* Paris : InterÉditions, 1980.

EVANS, C. *The Making of the Micro, A History of the Computer.* New York : Van Nostrand Reinhold, 1981.

FACKLAM, M. et FACKLAM, M. *The Brain.* New York : Harcourt Brace Jovanovich, 1982.

GARDNER, H. *Frames of Mind.* New York : Basic Books, 1983.

GOULD, S. J. *La mal-mesure de l'homme.* Paris : Ramsay, 1983.

HOFSTADTER, D. *Gödel, Escher, Bach.* Paris : InterÉditions, 1985.

JASTROW, R. *Au-delà du cerveau : de l'intelligence biologique à l'intelligence artificielle.* Paris : Hachette-Pluriel, 1984.

JEANNEROD, M. *Le cerveau machine.* Paris : Fayard, 1983.

JONES, E. *La vie et l'œuvre de Sigmund Freud 1856-1939* (6 vol.). Paris : P.U.F., 1969-1972.

ORNSTEIN, R. ; THOMPSON, R. et MACAULAY, D. *L'incroyable aventure du cerveau.* Paris : InterÉditions, 1986.

SANDERS, D. H. *Computers Today.* New York : McGraw-Hill, 1983.

SIMONS, G. *L'ordinateur est-il vivant ?* Paris : Londreys, 1984.

THOMPSON, C. *La psychanalyse, son évolution, ses développements.* Paris : Gallimard, 1956.

Appendice Le rôle des mathématiques

BELL, E. *La mathématique, reine et servante des sciences.* Paris : Payot, 1952.

Collectif. *Les progrès des mathématiques.* Paris : Belin/Bibliothèque pour la Science, 1981.

DANTZIG, T. *Le nombre : langage de la science.* Paris : Blanchard, 1974.

DAVIS, P.-J. et HERSH, R. *L'univers mathématique.* Paris : Gauthier-Villars, 1985.

FÉLIX, L. *L'aspect moderne des mathématiques.* Paris : Blanchard, 1957.

HILDEBRANDT, S. et TROMBA, A. *Formes optimales et mathématiques.* Paris : Belin, 1986.

HOGBEN, L. *Les mathématiques pour tous.* Paris : Payot, 1962.

RADEMACHER, H. et TOEPLITZ, O. *Plaisirs des mathématiques.* Paris : Dunod, 1967.

ROZSA, P. *Jeux avec l'infini, voyage à travers les mathématiques.* Paris : Le Seuil/Points Science, 1977.

SOLOMON, C. *Les mathématiques.* Paris : Larousse/Poche Couleur, 1970.

STEIN, S. K. *Les mathématiques, ce monde que créa l'homme.* Paris : Dunod, 1967.

VALENS, E. G. *The Number of Things.* New York : Dutton, 1964.

Index des noms

Index des sujets

Achevé d'imprimer
par Maury-Imprimeur-S.A.
45330 Malesherbes

Dépôt légal : Septembre 1986

LA VERSION FRANÇAISE

Elle a été assurée par une équipe de scientifiques universitaires spécialisés :

- *Pour la partie physique : Françoise Balibar, Claude Guthmann, Alain Laverne et Jean-Pierre Maury.*
- *Pour la partie biologie : Christine Coulondre, Marie-Jo Masse et John Reams.*

▸ reconnaître le don rare qui consiste à rendre compréhensible à l'esprit le moins scientifique le phénomène le plus ardu. En bon vulgarisateur, il ne sacrifie en rien la rigueur scientifique mais trouve le moyen d'émailler son propos d'anecdotes et d'images qui facilitent la compréhension tout en faisant de son texte une lecture aussi distrayante qu'instructive.

D'innombrables figures, schémas et planches viennent encore compléter l'information. Un appendice mathématique, une bibliographie et un index très fouillés font de cet ouvrage un véritable outil de travail pour ceux qui voudront l'avoir en permanence à portée de la main.

L'illustration de couverture représente la Géode, le bâtiment le plus spectaculaire de la Cité des Sciences et de l'Industrie de La Villette.
(Photo J.-P. Nacivet/Explorer)
Maquette de couverture : Tract.
(Photo Asimov D.R. traitée sur système Gixi).

ISAAC ASIMOV

L'univers de la science



Isaac Asimov est né en Union Soviétique. A l'âge de trois ans il arrive à New York avec ses parents immigrants qui ouvrent un petit commerce à Brooklyn. Enfant surdoué, il termine ses études secondaires à l'âge de quinze ans, puis étudie à Columbia University où il obtient son doctorat de chimie. De 1949 à 1958 il enseigne la biochimie à la faculté de médecine de Boston. Il y donne encore, à l'occasion, des conférences.

Asimov publie son premier livre, *Cailloux dans le ciel,* en 1950. Il devient rapidement le maître incontesté de la science-fiction, avec le cycle de *Fondation, Les robots, Le voyage fantastique,* etc. Sa production d'ouvrages de vulgarisation scientifique est également considérable, avec notamment, parmi ses œuvres les plus récentes, *La conquête du savoir.*

Mais Asimov n'en reste pas là. Il publie aussi des essais sur la Bible, des recueils de poèmes comiques, des ouvrages historiques, des essais sur Shakespeare, et même des romans policiers (*Une bouffée de mort,* entre autres).

Isaac Asimov, qui ne prend jamais l'avion, vit avec sa femme Janet Jeppson, écrivain elle aussi, dans un gratte-ciel de Manhattan. *L'univers de la science* est son 309e livre.

Réunie en un seul volume, c'est tout simplement la somme des connaissances scientifiques conquises par l'humanité, de l'Antiquité à nos jours, que nous offre cet auteur surprenant.

InterEditions
87 AVENUE DU MAINE 75014 PARIS

ISBN 2 7296 0141 4

